

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2003, 15:26:53 ; Search time 6842 Seconds
(without alignments)
11420.255 Million cell updates/sec

Title: US-10-019-633-1
Perfect score: 1910
Sequence: 1 ccacgctccgggtgtctcc.....aaaaaaaaaaaaaaaaaaaaa 1910

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEnbl.*

1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_on.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1910	100.0	1910	6	AX081182	AX081182 Sequence
2	1836	96.1	2573	6	AX081194	AX081194 Sequence
3	1238	64.8	1988	6	AX081200	AX081200 Sequence
4	1101.4	57.7	2782	6	AX081196	AX081196 Sequence
5	1043.4	54.6	2520	6	AX357082	AX357082 Sequence
6	1041.8	54.5	2551	8	AY133852	AY133852 Arabidops
7	1041.8	54.5	2847	8	AY056301	AY056301 Arabidops
8	1041.8	54.5	2880	6	AX015734	AX015734 Sequence
9	1028	53.8	1803	6	AX507020	AX507020 Sequence
10	962.4	50.4	2482	6	AX081198	AX081198 Sequence
11	959	50.2	1816	6	AX081184	AX081184 Sequence
12	531.6	27.8	914	6	AX357098	AX357098 Sequence
13	498.6	26.1	42544	8	SPBC106	AL110295 S.pombe c
14	487.8	25.5	3667	10	BC049119	BC049119 Mus muscu
15	477.2	25.0	4761	8	SPADA1TCP	X91498 S.pombe ada
16	476.6	25.0	2732	5	BC044154	BC044154 Danio rer
17	475	24.9	71088	8	NCB7F18	AL389891 Neurospor
18	471.6	24.7	2857	9	HSU16270	U16270 Human AMP d
19	471.6	24.7	3359	9	S47833	S47833 Homo sapien
20	471.6	24.7	3386	9	HUMAMPD2	M91029 Human AMP d
21	471.6	24.7	3702	9	BC007711	BC007711 Homo sapi
22	470	24.6	3745	9	AX025706	AK025706 Homo sapi
23	465.4	24.4	34506	8	SC8337	Z46659 S.cerevisia
24	463.8	24.3	3354	8	YSCAMD	M30449 Yeast (S.ce
25	461.8	24.2	3237	3	AY051429	AY051429 Drosophil
26	458.8	24.0	2341	9	HUMAMPD1	M60092 Human myoad
27	448.4	23.5	2391	9	AK097077	AK097077 Homo sapi
28	447.8	23.4	2241	6	AX457110	AX457110 Sequence
29	441.2	23.1	2275	10	RATAMPDA	J02811 Rattus norv
30	439.6	23.0	4098	10	BC040366	BC040366 Mus muscu
31	438.8	23.0	2757	3	AF238311	AF238311 Dictyoste
32	436.2	22.8	3490	10	D85596	D85596 Mouse mRNa
33	422	22.1	3915	6	AX683214	AX683214 Sequence
34	422	22.1	3915	10	RNU90888	U90888 Rattus norv
35	416	21.8	3680	9	HUMAMPD3B	M84721 Human AMP d
36	416	21.8	3732	9	HUMAMPD3	D12775 Homo sapien
37	411.8	21.6	2945	10	BC007183	BC007183 Mus muscu
38	330.4	17.3	1928	10	BC016662	BC016662 Mus muscu
39	308.8	16.2	45189	3	U46753	U46753 Caenorhadi
40	305.4	16.0	2212	9	HUMAMPD3C	M84722 Human AMP d
41	256.4	13.4	35870	3	AC018749	AC018749 Leishmani
42	256.4	13.4	40020	3	AC068783	AC068783 Leishmani
43	234	12.3	44319	2	AC013191	AC013191 Drosophil
44	234	12.3	92969	2	AC010844	AC010844 Drosophil
45	234	12.3	163958	3	AC010705	AC010705 Drosophil

ALIGNMENTS

RESULT 1
AX081182
LOCUS AX081182
DEFINITION Sequence 1 from Patent WO0109305.
ACCESSION AX081182
VERSION AX081182.1 GI:13170066
KEYWORDS Zea mays
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1
AUTHORS Caspar,T., Falco,S.C., Sakai,H., Weng,Z. and Hu,X.
TITLE Purine metabolism genes in plants

JOURNAL	Patent: WO 0109305-A 1 08-FEB-2001; E.I. DU PONT DE NEMOURS AND COMPANY (US) ; PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES	Location/Qualifiers
source	1..1910 /organism="Zea mays" /mol_type="genomic DNA" /db_xref="taxon:4577"
BASE COUNT	596 a 391 c 383 g 540 t
ORIGIN	
Query Match	100.0%; Score 1910; DB 6; Length 1910;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 1910; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 CCACGCGTCCGGGTGGCTCCATGGAGAGAGAGGTCATAAATGACCCCTGTACTCCAAA 60
Db	1 CCACGCGTCCGGGTGGCTCCATGGAGAGAGAGGTCATAAATGACCCCTGTACTCCAAA 60
Qy	61 CCTAACCCCCAACCCGTTCCACTTATGTGCTGAAACCAAGTCAGAGCATGTTTCCAAACT 120
Db	61 CCTAACCCCCAACCCGTTCCACTTATGTGCTGAAACCAAGTCAGAGCATGTTTCCAAACT 120
Qy	121 GTTGATGGCGTTATCCATGTTTATGCGGATAAAGATTGTACGGAGAGCATTTATCCTGTG 180
Db	121 GTTGATGGCGTTATCCATGTTTATGCGGATAAAGATTGTACGGAGAGCATTTATCCTGTG 180
Qy	181 GCTGATGCTACAACTTCTTCCACTGACITGGCATTTATTTCTCCGAGTACGGCTCGAGG 240
Db	181 GCTGATGCTACAACTTCTTCCACTGACITGGCATTTATTTCTCCGAGTACGGCTCGAGG 240
Qy	241 AACACAAGAACTGTCTGCCATAATCGGTTAAATCTTTTGAGCATAAGTTTAAATTCAT 300
Db	241 AACACAAGAACTGTCTGCCATAATCGGTTAAATCTTTTGAGCATAAGTTTAAATTCAT 300
Qy	301 CTGATGTTAAATGCGGATAGGGAATTTCTTGCCCAAGAGAATGCCCCACATCGTGATTTT 360
Db	301 CTGATGTTAAATGCGGATAGGGAATTTCTTGCCCAAGAGAATGCCCCACATCGTGATTTT 360
Qy	361 TACAATGTCAGGAAGTTGACATCTGCTTCATCATTCAGCATGATGAATCAAAAACAT 420
Db	361 TACAATGTCAGGAAGTTGACATCTGCTTCATCATTCAGCATGATGAATCAAAAACAT 420
Qy	421 CTGTTGAGGTTTCATAAAATCCAAAATAAGAAAGAACCTGTAGAGGTGGTCATTTTCAGA 480
Db	421 CTGTTGAGGTTTCATAAAATCCAAAATAAGAAAGAACCTGTAGAGGTGGTCATTTTCAGA 480
Qy	481 GATGGTACTTATATGACTTTTAAAGAGGTTTTTGAAGCTTGGACCTTAACTGGGTATGAT 540
Db	481 GATGGTACTTATATGACTTTTAAAGAGGTTTTTGAAGCTTGGACCTTAACTGGGTATGAT 540
Qy	541 CTGAATGTTGATTTGCTAGATGTCATCGACACAAAGCACATTTTCATCGTTTTCACAAA 600
Db	541 CTGAATGTTGATTTGCTAGATGTCATCGACACAAAGCACATTTTCATCGTTTTCACAAA 600
Qy	601 TTCAATCTAAAATACAAATCCATGTGCCCAAAGTAGGCTCAGAGAAATTTTCTCTCAAAACAA 660
Db	601 TTCAATCTAAAATACAAATCCATGTGCCCAAAGTAGGCTCAGAGAAATTTTCTCTCAAAACAA 660
Qy	661 GATAATCTTATTCGAAGCCGTTTTCTTGCTGAGTTTGACAAAGCAAGTTTCTCTGACCTTT 720
Db	661 GATAATCTTATTCGAAGCCGTTTTCTTGCTGAGTTTGACAAAGCAAGTTTCTCTGACCTTT 720
Qy	721 TCTGCTAGCAAAATATCAGATGGCAGAAATATAGGATTTCAATCTACGGAGGAAACAGAGT 780
Db	721 TCTGCTAGCAAAATATCAGATGGCAGAAATATAGGATTTCAATCTACGGAGGAAACAGAGT 780
Qy	781 GAATGGGACCAACTTGGCAAGTTGGATAGTGAACAATGAATTCGACAGTGGAAATGTTGTC 840
Db	781 GAATGGGACCAACTTGGCAAGTTGGATAGTGAACAATGAATTCGACAGTGGAAATGTTGTC 840
Qy	841 TGGCTGGTTCAGATTCCACGCTTATATAATGTGTACAAGGAAATGGGTATCGTTACATCA 900

RESULT 2
AX081194

LOCUS AX081194 2573 bp DNA linear PAT 27-FEB-2001
DEFINITION Sequence 13 from Patent WO0109305.
ACCESSION AX081194
VERSION AX081194.1 GI:13170072
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1.
REFERENCE
AUTHORS Caspar, T., Falco, S.C., Sakai, H., Weng, Z. and Hu, X.
TITLE Purine metabolism genes in plants
JOURNAL Patent: WO 0109305-A 13 08-FEB-2001;
E.I. DU PONT DE NEMOURS AND COMPANY (US) ; PIONEER HI-BRED
INTERNATIONAL, INC. (US)
FEATURES
Location/Qualifiers
1..2573
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="taxon:4577"
BASE COUNT 716 a 523 c 588 g 694 t 52 others
ORIGIN
Query Match. 96.1%; Score 1836; DB 6; Length 2573;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1836; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 12 GGTGCTCCATGGGAGAGGAGTCAATAAGCACCCCTGTACTCCAAACCTTAACCCCA 71
DB 738 GGTGCTCCATGGGAGAGGAGTCAATAAGCACCCCTGTACTCCAAACCTTAACCCCA 797
QY 72 CCCGTTCACTATGTGCGCTGAACCAAGTCAAGAGCATGTTTCCAAACCTGTTGAGCGGT 131
DB 798 CCCGTTCACTATGTGCGCTGAACCAAGTCAAGAGCATGTTTCCAAACCTGTTGAGCGGT 857
QY 132 TATCCATGTTATGCGGATAAAGATTGACGAGAGCATTTATCCTGCGGTGATGCTAC 191
DB 858 TATCCATGTTATGCGGATAAAGATTGACGAGAGCATTTATCCTGCGGTGATGCTAC 917
QY 192 AACCTTCTCACTGACTTGCATTATATCTCCGAGTAACGGCTGAGGAGACACAAAGAC 251
DB 918 AACCTTCTCACTGACTTGCATTATATCTCCGAGTAACGGCTGAGGAGACACAAAGAC 977
QY 252 TGTCTGCCAATAATCGGTTAAATCTTCTTGAGCATAAGTTTAAATCCATCTGATGTTAAA 311
DB 978 TGTCTGCCAATAATCGGTTAAATCTTCTTGAGCATAAGTTTAAATCCATCTGATGTTAAA 1037
QY 312 TCGGATAGGGAATTTCTTGCCAGAGAACTGCCCCACATCGTGATTTTACAATGTGAG 371
DB 1038 TCGGATAGGGAATTTCTTGCCAGAGAACTGCCCCACATCGTGATTTTACAATGTGAG 1097
QY 372 GAAGGTTGACACTCAATGTTTCATCTTACGATGATGATCAATCAAAACATCTGTTGAGGTT 431
DB 1098 GAAGGTTGACACTCAATGTTTCATCTTACGATGATGATCAATCAAAACATCTGTTGAGGTT 1157
QY 432 CATAAATCCAACTAAGAAAGAAACCTGATGAGGTGGTCAATTTTCAGAGATGGTACTTA 491
DB 1158 CATAAATCCAACTAAGAAAGAAACCTGATGAGGTGGTCAATTTTCAGAGATGGTACTTA 1217
QY 492 TATGACTTTAAAGAGGTTTTTGGAGGCTTGAAGCTTAACTGGGTATGATGAAATGTTGA 551
DB 1218 TATGACTTTAAAGAGGTTTTTGGAGGCTTGAAGCTTAACTGGGTATGATGAAATGTTGA 1277
QY 552 TTTGCTAGATGCTCCATGAGCAAAAGACATCTTCAATGTTTGGCAATTCATCTAAA 611
DB 1278 TTTGCTAGATGCTCCATGAGCAAAAGACATCTTCAATGTTTGGCAATTCATCTAAA 1337
QY 612 ATACAATCCATGTGGCCAAAGTAGGCTCAGAGAAATTTTCTCCTCAAAAGATAATCTTAT 671
DB 1338 ATACAATCCATGTGGCCAAAGTAGGCTCAGAGAAATTTTCTCCTCAAAAGATAATCTTAT 1397
QY 672 TCAAGGCGGTTTTCTTGTGATGAGTTGACAAAGCAAGTTTTCTCTGACCTTTCTGCTAGCAA 731

Db 1398 TCAAGGCGGTTTTCTTGTGATGAGTTGACAAAGCAAGTTTTCTCTGACCTTTCTGCTAGCAA 1457
QY 732 ATATCAGATGCGAGAAATATAGGATTTCAATCTACGGAAGGAAACAGAGTGAATGGGACCA 791
DB 1458 ATATCAGATGCGAGAAATATAGGATTTCAATCTACGGAAGGAAACAGAGTGAATGGGACCA 1517
QY 792 ACTTGCAAGTTGGATATGAGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 851
DB 1518 ACTTGCAAGTTGGATATGAGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 1577
QY 852 GATTCCAGCTTATATATGATGTAACAGAAATGGGTATCGTTACATCATTTCCAAATCT 911
DB 1578 GATTCCAGCTTATATATGATGTAACAGAAATGGGTATCGTTACATCATTTCCAAATCT 1637
QY 912 TCTTGACAAACATTTCTGTTCTCTTTTGAGGTTACTATTGATCAGCTTCAACCCACA 971
DB 1638 TCTTGACAAACATTTCTGTTCTCTTTTGAGGTTACTATTGATCAGCTTCAACCCACA 1697
QY 972 GCTCCATGTTCTGTAAGCAGGTTGTAGGTTGGACCTGTTGATGATGAAAGTAAACC 1031
DB 1698 GCTCCATGTTCTGTAAGCAGGTTGTAGGTTGGACCTGTTGATGATGAAAGTAAACC 1757
QY 1032 AGAAGGCGTCCAAAGAGCAGATGCCACACCTGAAACAGTGGACCAATGTTCAACCC 1091
DB 1758 AGAAGGCGTCCAAAGAGCAGATGCCACACCTGAAACAGTGGACCAATGTTCAACCC 1817
QY 1092 TGCATTTTCAATATATGCTACTACTGCTATGCTTAACTTATTCACCTTAAACAGCTCG 1151
DB 1818 TGCATTTTCAATATATGCTACTACTGCTATGCTTAACTTATTCACCTTAAACAGCTCG 1877
QY 1152 TGAGTCAAGGGAATGACCACTATCAAAATTCCTGCCACATGCTGGAGGCTGGAGATGT 1211
DB 1878 TGAGTCAAGGGAATGACCACTATCAAAATTCCTGCCACATGCTGGAGGCTGGAGATGT 1937
QY 1212 TGATCATTGGCAGGACATTTCTTCTGTCACAAATATACATGGAATTAATCTAAG 1271
DB 1938 TGATCATTGGCAGGACATTTCTTCTGTCACAAATATACATGGAATTAATCTAAG 1997
QY 1272 GAAGTCTCTGCTGCTCAGTACTTGTACTTCTTGGTCAGATGCTGCGGATGCTCC 1331
DB 1998 GAAGTCTCTGCTGCTCAGTACTTGTACTTCTTGGTCAGATGCTGCGGATGCTCC 2057
QY 1332 ATTGAGCAACAACTCTTATTTCTTGACTATCATCGCAACCTTTTCCAAAGTCTTCTCA 1391
DB 2058 ATTGAGCAACAACTCTTATTTCTTGACTATCATCGCAACCTTTTCCAAAGTCTTCTCA 2117
QY 1392 ACAGGCTCTGAATGCTCTATTTATCTACGGATGACCTTTTGCAAAATTCACCTGACAAAGA 1451
DB 2118 ACAGGCTCTGAATGCTCTATTTATCTACGGATGACCTTTTGCAAAATTCACCTGACAAAGA 2177
QY 1452 ACCATTGCTGGAAGAAATACAGCATTTGCTTCTGCTTGGAGCTCAGTTCTTGTGATTT 1511
DB 2178 ACCATTGCTGGAAGAAATACAGCATTTGCTTCTGCTTGGAGCTCAGTTCTTGTGATTT 2237
QY 1512 ATGCGAAATTCGAGGAACTCTGTTTACCAATCTGGGTTTTTCAATGCTCTCAAGGCGCA 1571
DB 2238 ATGCGAAATTCGAGGAACTCTGTTTACCAATCTGGGTTTTTCAATGCTCTCAAGGCGCA 2297
QY 1572 CTGGAATTTGAAGAACTCTTTCAAAAGAGGACCTGCTGGAAATGATATTCAGAAACCA 1631
DB 2298 CTGGAATTTGAAGAACTCTTTCAAAAGAGGACCTGCTGGAAATGATATTCAGAAACCA 2357
QY 1632 TGTAACCCCATCAGGTTTCAATTTTAGAGAGATCATCTGGGAATGAAATGAAATGAAATGAA 1691
DB 2358 TGTAACCCCATCAGGTTTCAATTTTAGAGAGATCATCTGGGAATGAAATGAAATGAAATGAA 2417
QY 1692 GTACTCTGACAAATGAGATCTTAAATACAGAGAGCTGGACCTGTAAGATGCTCCAGCTCG 1751
DB 2418 GTACTCTGACAAATGAGATCTTAAATACAGAGAGCTGGACCTGTAAGATGCTCCAGCTCG 2477
QY 1752 TGTATACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1811

Db 1700 GG 1701

RESULT 4
AX081196
LOCUS AX081196
DEFINITION Sequence 15 from Patent WO0109305.
ACCESSION AX081196
VERSION AX081196.1 GI:13170073
KEYWORDS
SOURCE
ORGANISM
Oryza sativa
Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartioideae; Oryzeae; Oryza.

REFERENCE
1
AUTHORS Caspar, T., Falco, S.C., Sakai, H., Weng, Z. and Hu, X.
TITLE Purine metabolism genes in plants
JOURNAL Patent: WO 0109305-A 15 08-FEB-2001;
E.I. DU PONT DE NEMOURS AND COMPANY (US); PIONEER HI-BRED
INTERNATIONAL, INC. (US)
FEATURES
Location/Qualifiers
1..2782
/organism="Oryza sativa"
/mol_type="genomic DNA"
/db_xref="taxon:4530"
BASE COUNT 798 a 640 c 591 g 753 t
ORIGIN

Query Match 57.7%; Score 1101.4; DB 6; Length 2782;
Best Local Similarity 78.5%; Pred. No. 3.7e-219;
Matches 1318; Conservative 0; Mismatches 361; Indels 0; Gaps 0;

QY 13 GTTGCTCCATGGAGAGGAGCTAATAGTCCCTGCTACTCCAAACCTAACCCAC 72
DB 768 GTTGCTCCATGGAGAGGAGGAACTAATGCTAGTCTAGTCTCCAAACCTAAC 827

QY 73 CGGTTCACCTATGTGCTGAACCAAGTCAGAGCATGTTTCCAAACCTGTGATGCGTT 132
DB 828 CTTTCTATTACAGCAGCAGACTAAACTGAACATCTTTTGAATGTTGATGTT 887

QY 133 ATCCATGTTTATGCGGATTAAGATTGTACGGAGAGCATTTATCTGTGGCTGATGTACA 192
DB 888 ATTCATGTATACCCCAATAAGACGCTAAAGAAAGAACTATCTCTGTGCTGATGCTACT 947

QY 193 ACCTCTTTCACCTGCTGATTTATATCTCCAGTAAACGCTGCGAGGAGACACAGACT 252
DB 948 ACCTTTTTCACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1007

QY 253 GTCTGCCATTAATCGGTAAATCTTCTGAGCATTAAGTTTAAATTCATCTGATGTTAAAT 312
DB 1008 GTATGTTTAAACGTTTAAATCTTCTAGAACAGAAATCAATCTTCAATTTGATGCTCAAT 1067

QY 313 GCGGATAGGAAATTTCTGCCAGAAAGACTGCCCAACATCGTGATTTTAAATGTCAGG 372
DB 1068 GCGGATAGGAAATCTTCTGCTCAGAAAGCTGCACCCCATCGGGAATCTTCAATGTCAGG 1127

QY 373 RAGGTTGACATCATGCTTCAATTCATTCAGCATGCTGAATCAAAACATCTGTTGAGTTTC 432
DB 1128 AAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1187

QY 433 ATAAATCCAAACTAAGAAAGAACTGATGAGGTGCTATTTTCAGAGATGCTACTTAT 492
DB 1188 ATCAAGTCCAACTGAGGAAAGAACTGACGAGGTGATTTTGTAGATGCTACTTAT 1247

QY 493 ATGACTTTAAAGAGGTTTTCAGAGCTTGGACTTAACCTGAGTATGATCTGAATGTTGAT 552
DB 1248 TTGACTCTTAAGAGGTTTTCAGAGCTTGGAGCTTGAATGATGATGATGATGATGAT 1307

QY 553 TTGCTAGATGCTCAGCAGAAAGCAGCATTTTCATCTGTTTTCAGAAATTCATCTAAA 612
DB 1308 CTTCTAGATGCTCAGCAGAAAGCAGCATTTTCATCTGTTTTCAGAAATTCATCTAAA 1367

QY 613 TACAATCCATGTGGCCAAAGTAGGCTCAGAGAAATTTTCTCTCAAACAAGATAATCTTATT 672
DB 1368 TATAATCTTGTGGCCAAATCCCGGCTGAGGAGATCTTTCTTAAACAGGACAACTTATT 1427

QY 673 CAAGCCGCTTTCTGTGAGTTGACAAAGCAAGTTTCTCTGACCTTCTGTGACGAAA 732
DB 1428 CAAGCCGCTTTCTGTGAGTTGACAAAGCAAGTTTCTCTGATCTTTGAAAGCAAGTAAA 1487

QY 733 TATCAGATGGCAGAAATATAGGATTTCAATCTACGGAAGGAAACAGAGTGAATGGACCAA 792
DB 1488 TATCAGATGGCTAGTATAGATATCTATCTATGGGAGAAAGATGATGGATCAG 1547

QY 793 CTTGCAAGTTGGATAGTGAACAATGAATTCACAGTGGAAATGTTGCTGGCTGGTTTCA 852
DB 1548 ATGGCAAGCTGGATAGTGAATAATGAATTTGACAGCGAGATGTTGTTGGTTAATTCAG 1607

QY 853 ATTCACGCTTATATATATATGTTGACAGAAATGGGTATCGTTACATCAATTCAAAATCTT 912
DB 1608 ATTCCTCGGATATACAAATGTATACAGGAGATGGAAACAATCAATTTCTTCCAGAACCTC 1667

QY 913 CTTGACAACTTTTCTGTTTCTCTTTTCTGAGGTTTACTATTGATCCAGCTTTCACACCCACAG 972
DB 1668 CTTGACAACTTTTCTGCTCTTTTCTGAGTAACTGTTGATCTGCTTCACTCCTCAG 1727

QY 973 CTCCATGCTTCTCTGAAGCAGGTTGTAGGTTTGGACCTGGTTGATGATGAAGTAAACCA 1032
DB 1728 CTCCATGTTTCTTGGCAACAGTCGTTGGGCTGGATTTAGTGGATGATGAAGCAACCA 1787

QY 1033 GAAAGGCTCCAAACAAAGCAGATGCCACACCTGAAACAGTGGACCAATGTGTTCAACCTT 1092
DB 1788 GAGAGACCCCAACAAACACATGCTACACCTGAGCAATGGAATTAATGTTTCAATCCA 1847

QY 1093 GCATTTTCATATATGCTGCTACTGCTGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1152
DB 1848 GCATGATGATATATGCTGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1907

QY 1153 GAGTCAAAAGGGAATGACCACTATCAAAATTCGCTCCACATGCTGGAGAGGCTGGAGATGTT 1212
DB 1908 GAGTCAAAAGGATGACCAACCAATCAAACTTCGCTCCACATGCTGGGAGGCTGGAGATTT 1967

QY 1213 GATCACTGGCAGGAGCAATTTCTCTCTGTGACAAACATATCACATGGAATTAATCTAAG 1272
DB 1968 GATCATCTGCTGAGCAGATTTCTTACTTCTCATATATTTGCTCAGCGGCTTAAATTTAAG 2027

QY 1273 AAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1332
DB 2028 AAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2087

QY 1333 TTGAGCAACAACTCTCTTATTTTCTGATCATCGCAACCCCTTTTCCAACTGTTCTTCCAA 1392
DB 2088 TTGAGCAACAACTCATTTGTTTATTTGATTTATCAACGAAACCTTTTCCCAACATTTTCTCTA 2147

QY 1393 CGAGTCTGAATGCTCATATCTACGATGACCCCTTTTGCATAATTCACCTGACAAAAGAA 1452
DB 2148 AGAGGCTTAAAGCTTTCTCTATCAACCCGATGACCCCTTTTGCATAATTCACCTGACAAAAGAA 2207

QY 1453 CCATTTGTTGGAAGATACAGCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1512
DB 2208 CTTTGGTTGGAAGATATAGCATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2267

QY 1513 TCGCAAAATTCGAGGAACTCTGTTTACCAATCTGGGTTTTTCCATGCTCTCAAGCGGCAC 1572
DB 2268 TGTGAATTTGCTGGAATTTCTGTGTACAGTCTGTTTCTCTCATAGCTCAAGTCAAC 2327

QY 1573 TGGATTTGTAAGAACTACTTCAAAGAGGACCTGCTGGAATGATATTCAGAGAACCAAT 1632
DB 2328 TGGATTTGGAAGAACTACTTCAAAGAGGCTCATGAGCAATGACATTCACAGCAAAAT 2387

QY 1633 GTACGACATCAGGTTTCAATTTAGAGATGATCTGGAGAAATGAATGAATGAATGAAT 1691
DB 2388 GTTCTCATCAGGATTTGAATTCGACACTATTTGGAAAGAAAGAAATGGAGCTAAT 2446

```

RESULT 5
AX357082
LOCUS AX357082 2520 bp DNA linear PAT 13-FEB-2002
DEFINITION Sequence 1 from Patent WO0206319.
ACCESSION AX357082
VERSION AX357082.1 GI:18674263
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eucosids II; Brassicales; Brassicaceae; Arabidopsis.
1
REFERENCE
AUTHORS Schulz, A.; Streiber, W.; Hanke, C.; Schmidt, F. and Schubel, A.
TITLE Nucleic acid molecule which codes for a plant amp deaminase
JOURNAL Patent: WO 0206319-A 1 24-JAN-2002;
Aventis CropScience GmbH (DE)
FEATURES
source
1..2520
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
1..2520
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD23276.1"
/db_xref="GI:18674264"
/translation="MEPNYQLALALFASVAVSGFPMHFKALNLVLKERKERN
PGDBEPQNTLVRSSQVRKQNDQYRSPALPDATPFTDGGGGGGDTGNSGHVY
VDEIFPLRLTPSEGASVHGASIRKTSFVRPIPSKSPVASASAFESVEESDD
DNLNTEGSDASVLOANGDNEMPADNEBQIISMAASSMRSHSVSDLHGVDPIAA
DILRKEPEOTFRLNVLEVTDSDEAVKCLQECLELRKRYVFOETVAPKEKVIS
DPSTKPNTPEFANHYPOGSDHCFENQDGVHVFANKADELFPVADATATFDLHH
VLKVIAGNIRTLCHRLQLLEQFNLHMLNADREFLAQKSPHPRDFYNVRKVDTHV
HHSACNQKHLRFITKSLRKQEPDEVIIFRDGYTLTREVFSLEDLTGVDLNDVLDV
HADSTFRHDFLKNYPCGGRLEIFLQDNLTQGRFLGEITKQVFSDELSKYQ
MAEYRISIVGRKMSWDOLASWINDLSYVWMLIQLPRLNYIKDMGIVTSFQNI
LONIFLPEATVPDPSQLHVLKOVVGLVDDESKPERRPKHMTPTPAQWNAF
NPASVYVYCYANLVKLKRESKGMTITILRPHSGEAGDIDLAAFLTCHSIAG
INRKSVPVQYLYLQIAMSPLSNLSFLDYHRNPPFPFLKLNLSLSTDPLQ
IHLEKPELVYESIAASVWKLSDCLCEIARNSVYQSGFSLKSHWIGKDYKRGPD
GNDIHKTNYPHIRVEFRDITWKEEMQVYLGKAVISDEVVP"
BASE COUNT 699 a 541 c 572 g 708 t
ORIGIN
Query Match 54.8%; Score 1043.4; DB 6; Length 2520;
Best Local Similarity 76.0%; Pred. No. 4.5e-207;
Matches 1287; Conservative 0; Mismatches 406; Indels 0; Gaps 0;
QY 13 GTTGCTCCATGGGAGAGGAGTCAATAATGACCCCTGACTCCAAAACCTAACCCCAAC 72
D 13 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 799 GTTGACACCATGGGAAAAGAGTCATATCTGATCTCTAGTACTCCAAAGCCTTAATACAGAG 858
QY 73 CGGTTCACATATGTGCTGAACCAAAAGTCAGAGCATGTTTTCACAACTGTGATGCGGTT 132
D 73 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 859 CCATTTGCACATCTCTCAGGGAATCTGATCATTTGTTTGGAGATGCAAGATGGGTT 918
QY 133 ATCCATGTTATCGCGATAAAGATTGTACGGAGAGCATTTATCTCTGGCTGATGCTACA 192
D 133 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 919 GTCCACGTGTTTGCAAAATAAGATGCAAAAAGAGATCTCTTCCCGGTAGCTGATGCCACA 978
QY 193 ACCTCTTCACCTGATGATATATTTCTCCGAGTAACGGCTGCAGGGAACACAAACT 252
D 193 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 979 GCGTTTTTCACTGATGATCACTGACTCAAGTCAAAAGTCAATAGCTCAGGAAACATCCCGGACT 1038
QY 253 GTCTGCATAATCGGTAAATCTTCTTGAGCATAAAGTTTAAATTCATCTGATGTTAAAT 312
D 253 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 1039 TTGTGCCACCGTCTGACTAGTCTCTAGAACAGAAATTAATCTCCATTTGATGCTTAAT 1098
QY 313 GCGGATAGGAATTTCTTGCCCAAGAGACTGCCCAACATGCTGATTTTACAATGTCAGG 372
D 313 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 1099 GCGGATAAAGAAATTTCTTGCTCAAAAAGATGCACCATGCTGATTTTATACGTTAGG 1158

```

```

QY 373 AAGGTTGACACTCATGTTTCATCATTCAGCATGCAATCAAAAAATCATCTGTTGAGGTTTC 432
D 373 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 1159 AAGGTCGACACTCATGTCATCATTCAGCTGTCATGACCAAGAACACCTTTTAAGGTTT 1218
QY 433 ATAAATCAAACTAAGAAAAGAACTGATGAGGGTGCATTTTTCAGAGATGGTACTTAT 492
D 433 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 1219 ATTAAGTCAAAAGCTCCGAAAAGAACCCGATGAGGTAGTAAATATTCGAGATGGAACATAT 1278
QY 493 ATGACTTTAAAGAGGTTTTCAGAGCTTGAGACTTAACCTGGGTATGATCTGAATCTTAT 552
D 493 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 1279 TTGACCTTTGAGAGAAAGTTTTTTCAGAGCTTGAGTCTGAGTGGGTATGACTGAACTGCAC 1338
QY 553 TTGCTAGATGTCATGACAGAAAAGACATTTTCATCGTTTTCAGAAAATTCATCTAAA 612
D 553 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 1339 CTTTGGATGTTTCAGAGCAAAAAGTACCTTTTCATCGTTTTCAGTAAGTTCAACCTTAAG 1398
QY 613 TACAATCCATGTGGCCAAAGTAGGCTCAGAGAAAATTTTCTCAAAACAAGATAATCTTAT 672
D 613 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 1399 TATAACCTTTGTTGGTCAAAAGTAGGCTTAGGAGATTTTCTTAAACAGGATAATCTCATC 1458
QY 673 CAAGCCGTTTTCCTGCTGAGTTGACAAAGCAAGTTTTCTCTGACCTTTCTGCTAGCAAA 732
D 673 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 1459 CAAGGTCGATTTCTTGGTGAGATAACAAAGCAAGCTTTCTCTGACCTTTGAAGCTAGTAAA 1518
QY 733 TATCAGATGGCAGAAATATAGGATTTCAATCTACGGAAGAAACACAGAGTGAATGGGACCAA 792
D 733 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 1519 TATCAGATGGCTGATACAGATATCTATATATGSCAGAAAATAGCGAGTGGGACCAA 1578
QY 793 CTTGCAAGTTGGATAGTGAACAATGAATGCAAGTGAATGTTTGTCTGGCTGGTTTCAG 852
D 793 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 1579 CTCGCTAGTTGGATTTGAAACAATGATCTATACAGTGAATGTTTGTCTGGTTAAATTCAG 1638
QY 853 ATTCACCGTTTATATAATGTGTACAGAAATGGGTATCGTTACATCATTCCTCAAAATCTT 912
D 853 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 1639 CTCACACGCTTGTACAAATTTTACAAGACATGGGTATTTGACATCGTTTCCAGAAATATC 1698
QY 913 CTTGACAAATTTTCGTTTCTCTTTTTCAGGGTTACTATTTGATCCAGCTTTCACACCCACAG 972
D 913 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 1699 CTGGACATATATTCATTTCTCTGTTTGAAGCCAGGTAGATCTCTGATTTCCATCCTCAG 1758
QY 973 CTCATGTTCTCTCAAGCAGGTTGTAGGGTTGGACCTGGTTGATGATGAAGATGAACCA 1032
D 973 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 1759 CTCATGTTTTCGAGCAGGTTGTGATTTGATTTGTTGTTGATGATGAAGAACCAACCT 1818
QY 1033 GAAAGGCTCCAAAGCAATGCCCACTGACCACTGACCACTGACCACTGTTGTTCAACCT 1092
D 1033 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 1819 GAAAGACGTCACAAACACATGCCCACTCCAGCTCAATGGACTTAACGCAATTCATCTCT 1878
QY 1093 GCATTTTCATATTTATGCGTACTCTGCTATGCTAACTTTATTCACCTTAAACAAAGCTGCGT 1152
D 1093 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 1879 GCATTTTCGATATGTTCTATCTATTTGTTATGCTTAACCTCTATGTTTAAATGAAGCTTCA 1938
QY 1153 GAGTCAAAAGGGAATGACCACTATCAAAATTCCTGTCACATGCTGGAGAGGCTGGAGATGTT 1212
D 1153 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 1939 GAGTCAAAAGGGAATGACTACTATACGCTACGACCACTTCTGGAGAGGCTGGTGCATTT 1998
QY 1213 GATCACTTTGGCAGCAGATTTCTCTCTGTCAACATATCACTGGAATTAATCTAAGG 1272
D 1213 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 1999 GACCACTTTGGCTGCTACGTTTCTAACTAGCCATAGCACTGGAATTAATCTGCGCA 2058
QY 1273 AAGTCTCTGCTGCTTTCAGTACTTGTACTATCTTGTGTCAGATGTTGCTGGCGATGTCGCCA 1332
D 1273 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 2059 AAGTCTCTGCTGCTTTCAGTACTTGTACTCTCTGCGCCAGATTTGCTTGGCCATGTCACCA 2118
QY 1333 TTGAGCAACAACCTCTTATTTCTTCACTATCATCAACACCTTTTCCAAAGTTCCTTCCAA 1392
D 1333 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 2119 CTGAGCAACAACCTCTTGTGTTCTAGATTACCAACCGGAACCCGTTTCTGTTGTTTCTTCA 2178
QY 1393 CGAGGTCGATGCTCTCATTTCTAGGATGACCTTTTGCAAAATTCACCTGCAACAAAGAA 1452
D 1393 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 2179 AGAGGTCGATGTTTCTGCTACTGATGACCCCTTCAGATTCACCTTAACCTAAGAA 2238

```

QY 1453 CCATTGCGGAAGATACAGCAATGCTGCTCGCTGTGGAAGCTCAGTTCGTGATTTA 1512
 Db 2239 CTTCTCGTGAAGAGATAGCATAGCTGCATCAGTATGGAAGCTGAGTGGTGCACCTG 2298
 QY 1513 TCGCAAAATGCGAGGAATCTGTTTACCAATCTGGGTTTTCACATGCTCTCAAGCGCAC 1572
 Db 2299 TCGGAGATAGCTCGTAACCTCAGTGTACCACTCAGTTCCTCACGCCCTGAGTGCAC 2358
 QY 1573 TGGATTGTAAGAACTACTTTCAAAAGAGAGCACTGCTGGAAATGATATTACAGAACCAAT 1632
 Db 2359 TGGATTGGAAGAAATTAATCTACAAAGAGAGCACTGATGGAAGAGCACTTACCAAAACAAAC 2418
 QY 1633 GTACGCGCATCAGAGGTTCAATTTAGAGAGATGATCTGGAGAAATGAATGAATGATGTG 1692
 Db 2419 GTGCCACACATAAGGGTGGAGTTCCTGTGACAGATCTGGAAGAGGAGATGCAACAGTT 2478
 QY 1693 TACTCTGACAATG 1705
 Db 2479 TATCTGGGCAGG 2491

RESULT 6
 LOCUS AY133852 2551 bp mRNA linear PLN 18-SEP-2002
 DEFINITION Arabidopsis thaliana clone U10978 putative AMP deaminase
 (At2g38280) mRNA, complete cds.
 ACCESSION AY133852
 VERSION AY133852.1 GI:22136883
 KEYWORDS FLI CDNA
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1 (bases 1 to 2551)
 Yanada K., Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hsuan V.W.,
 Lee J.M., Quach H.L., Tang C.C., Toriumi M., Wallender E.K., Wong C.,
 Wu H.C., Yu G., Yuan S., Bowser L., Carninci P., Chen H., Cheuk R.,
 Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
 Kawai J., Kim C., Lam B., Lin J., Miranda M., Narusaka M., Nguyen M.,
 Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A.,
 Shinozaki K., Davis R.W., Ecker J.R. and Theologis A.
 Arabidopsis Open Reading Frame (ORF) Clones
 Unpublished
 2 (bases 1 to 2551)
 Yanada K., Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hsuan V.W.,
 Lee J.M., Quach H.L., Tang C.C., Toriumi M., Wallender E.K., Wong C.,
 Wu H.C., Yu G., Yuan S., Bowser L., Carninci P., Chen H., Cheuk R.,
 Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
 Kawai J., Kim C., Lam B., Lin J., Miranda M., Narusaka M., Nguyen M.,
 Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A.,
 Shinozaki K., Davis R.W., Ecker J.R. and Theologis A.
 Direct Submission
 TITLE Submitted (17-JUN-2002) Plant Gene Expression Center, 800 Buchanan
 JOURNAL Street, Albany, CA 94710, USA
 REFERENCE The RIKEN Genomic Sciences Center (GSC) members carried out the
 collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN
 Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J.,
 Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
 Hayashizaki, Y. and Shinozaki, K.
 The Salk, Stanford, PGEC (SSP) Consortium members constructed and
 sequenced the pUNI (ORF) clones using the RAFL cDNAs: Yanada, K.,
 Chan, M.M., Chang C.H., Dale, J.M., Deng, J.M., Hsuan, V.W., Lee, J.M.,
 Quach H.L., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C.,
 Wu, H.C., Yu, G., Yuan, S., Bowser, L., Chen, H., Cheuk, R.,
 Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Miranda, M.,
 Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M.,
 Shinn, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and
 Theologis, A.

QY 1453 CCATTGCGGAAGATACAGCAATGCTGCTCGCTGTGGAAGCTCAGTTCGTGATTTA 1512
 Db 2239 CTTCTCGTGAAGAGATAGCATAGCTGCATCAGTATGGAAGCTGAGTGGTGCACCTG 2298
 QY 1513 TCGCAAAATGCGAGGAATCTGTTTACCAATCTGGGTTTTCACATGCTCTCAAGCGCAC 1572
 Db 2299 TCGGAGATAGCTCGTAACCTCAGTGTACCACTCAGTTCCTCACGCCCTGAGTGCAC 2358
 QY 1573 TGGATTGTAAGAACTACTTTCAAAAGAGAGCACTGCTGGAAATGATATTACAGAACCAAT 1632
 Db 2359 TGGATTGGAAGAAATTAATCTACAAAGAGAGCACTGATGGAAGAGCACTTACCAAAACAAAC 2418
 QY 1633 GTACGCGCATCAGAGGTTCAATTTAGAGAGATGATCTGGAGAAATGAATGAATGATGTG 1692
 Db 2419 GTGCCACACATAAGGGTGGAGTTCCTGTGACAGATCTGGAAGAGGAGATGCAACAGTT 2478
 QY 1693 TACTCTGACAATG 1705
 Db 2479 TATCTGGGCAGG 2491

FEATURES
 source
 1..2551
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="2"
 /clone="U10978"
 /note="This clone is in pUNI 51."
 ecotype: Columbia
 1..2551
 /gene="At2g38280"
 1..2520
 /gene="At2g38280"
 /codon_start=1
 /evidence=experimental
 /product="putative AMP deaminase"
 /protein_id="AA091786.1"
 /db_xref="GI:22136884"
 /translation="MEPIYCLALALPGLSFAVSGFPMHFKALNLVLRGKREKEN
 PDGEPQNPILRRRSQVRKNDQYGRSPASLPDPTFTDGGGGGGDGTGRSGHHVY
 VDETPGLPLRHTPTSEGRSVHGASSIRKTSFVRPISPKSPVASAFSEVESDDDD
 DNLNSELGLDASYLOANGDNEMPADNEEQISMAASSMIRSHSVGDLHGVPDPIAA
 DILRKEPEQETFVRLNVLPEVTSDEVEAYKLOECLERKRYVQETVAPEWEKEVIS
 DPSTPKNTEPFAHYPOGKSDHCPQEMODGVVHVFAKDKEDLFPADATATFTDLHH
 VLKVIAGNIRTLCRLVLEOKENLHMLNADKEFLAKSAAPHRDFYNNRVKVDTHV
 HHSACMNQKHLIRIKSLRKEPDEVVIFRDGTLYTLREVPESLDLGYDNLVDLLDV
 HADKSTFRHDKFNLYNCPGQSRLEIFLQDNLIQGRFLGELITTKQVDSLEAKYQ
 MAEYRISYGRKSEWDLASVIVNNDLYSENVMVLIQLPLNYIKDMPTVTSFQNI
 LDNIFILFEATVPDPSHPOLHFKQWGVFDLVDDSKPKERRTKMPTAOWTNAF
 NPAFYVYVYCYANLYVNLKRSKGWTTITLRPHSGEAGDIDHAAATFLTCHSIAHG
 INLRKSVLOYLYLAQIGLAMSPLSNNSFLDYHRNPFPVFLRGLNVSLSLTDPLQ
 IHLTKPELVEYSIAPASVWKLSCDLCLEARNSVYQSDFVSHALAKSHWIGDKYIRKRPD
 IDLHKTNVPHIRVEFRDFTWKEEMQVYLGKAVISDEVVP"
 2551
 /3'UTR
 2551
 /gene="At2g38280"
 713 a 548 c 576 g 714 t
 BASE COUNT
 ORIGIN
 Query Match 54.5%; Score 1041.8; DB 8; Length 2551;
 Best Local Similarity 76.0%; Pred. No. 9.6e-207;
 Matches 1286; Conservative 0; Mismatches 407; Indels 0; Gaps 0;
 QY 13 GTTCCTCATCGGAGAGAGGTCATAAATGACCCCTGTACTCCAAAACCTTAACCCCAAC 72
 Db 799 GTTCACCATCGGAGAGAGGTCATAAATGACCCCTGTACTCCAAAACCTTAACCCCAAC 858
 QY 73 CCGTTTCACTTATGTCCTGACCAAGTCAGAGCATGTTTCCAAACTGTTGTATGGCGCTT 132
 Db 859 CCATTTCGACACTATCTTCAGGGAATCTGATCACTGTTTGTAGATCAAGATGGGGTT 918
 QY 133 ATCCATGTTTATCGGATAAAGATTGTACGGAGAGCAATTTATCTGTGGCTGATGCTACA 192
 Db 919 GTCCACGCTGTTGCAATTAAGATGCAAAAGAACATCTTCCCGGTAGCTGATGCCACA 978
 QY 193 ACCTTCTTCACTGATGCTATATATTTCTCGAGTACGGCTGACGAGAACACAAAGAACT 252
 Db 979 GCGTTTTTCACTGATGCTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGT 1038
 QY 253 GTCTGCCATTAATCGTTTAAATCTTTTGTAGCATAGTTTAAATCTCATCTCATGTTAAAT 312
 Db 1039 TTGTGCCACCGTGCATGCTCTAGAACAGAAATTTATCTCCATTTGATGCTTAAT 1098
 QY 313 GCGGATAGGAATTTCTTGCCAGAGACTGCCCCACATCGTGATTTTTTACAATGTGTCAGG 372
 Db 1099 GCGGATAAAGAAATTTCTTGCTCAAAAAGTGCACCATCGTGATTTTTTATAACGTAGG 1158
 QY 373 AAGTTGACACTCATGTTTCATTCATTCAGCATGCAATGAATCAAAAACATCTGTTGAGGTT 432
 Db 1159 AAGTTCGACACTCATGTTGTCATTCATTCAGCTTGCATGAACAGAACACCTTTTAAAGGTTT 1218

/PGEC) contributed equally to this work as PIs.
 Annotation is based on the January 2002 version of the Arabidopsis
 genome submitted to GenBank.
 Location/Qualifiers
 1..2551
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="2"
 /clone="U10978"
 /note="This clone is in pUNI 51."
 ecotype: Columbia
 1..2551
 /gene="At2g38280"
 1..2520
 /gene="At2g38280"
 /codon_start=1
 /evidence=experimental
 /product="putative AMP deaminase"
 /protein_id="AA091786.1"
 /db_xref="GI:22136884"
 /translation="MEPIYCLALALPGLSFAVSGFPMHFKALNLVLRGKREKEN
 PDGEPQNPILRRRSQVRKNDQYGRSPASLPDPTFTDGGGGGGDGTGRSGHHVY
 VDETPGLPLRHTPTSEGRSVHGASSIRKTSFVRPISPKSPVASAFSEVESDDDD
 DNLNSELGLDASYLOANGDNEMPADNEEQISMAASSMIRSHSVGDLHGVPDPIAA
 DILRKEPEQETFVRLNVLPEVTSDEVEAYKLOECLERKRYVQETVAPEWEKEVIS
 DPSTPKNTEPFAHYPOGKSDHCPQEMODGVVHVFAKDKEDLFPADATATFTDLHH
 VLKVIAGNIRTLCRLVLEOKENLHMLNADKEFLAKSAAPHRDFYNNRVKVDTHV
 HHSACMNQKHLIRIKSLRKEPDEVVIFRDGTLYTLREVPESLDLGYDNLVDLLDV
 HADKSTFRHDKFNLYNCPGQSRLEIFLQDNLIQGRFLGELITTKQVDSLEAKYQ
 MAEYRISYGRKSEWDLASVIVNNDLYSENVMVLIQLPLNYIKDMPTVTSFQNI
 LDNIFILFEATVPDPSHPOLHFKQWGVFDLVDDSKPKERRTKMPTAOWTNAF
 NPAFYVYVYCYANLYVNLKRSKGWTTITLRPHSGEAGDIDHAAATFLTCHSIAHG
 INLRKSVLOYLYLAQIGLAMSPLSNNSFLDYHRNPFPVFLRGLNVSLSLTDPLQ
 IHLTKPELVEYSIAPASVWKLSCDLCLEARNSVYQSDFVSHALAKSHWIGDKYIRKRPD
 IDLHKTNVPHIRVEFRDFTWKEEMQVYLGKAVISDEVVP"
 2551
 /3'UTR
 2551
 /gene="At2g38280"
 713 a 548 c 576 g 714 t
 BASE COUNT
 ORIGIN
 Query Match 54.5%; Score 1041.8; DB 8; Length 2551;
 Best Local Similarity 76.0%; Pred. No. 9.6e-207;
 Matches 1286; Conservative 0; Mismatches 407; Indels 0; Gaps 0;
 QY 13 GTTCCTCATCGGAGAGAGGTCATAAATGACCCCTGTACTCCAAAACCTTAACCCCAAC 72
 Db 799 GTTCACCATCGGAGAGAGGTCATAAATGACCCCTGTACTCCAAAACCTTAACCCCAAC 858
 QY 73 CCGTTTCACTTATGTCCTGACCAAGTCAGAGCATGTTTCCAAACTGTTGTATGGCGCTT 132
 Db 859 CCATTTCGACACTATCTTCAGGGAATCTGATCACTGTTTGTAGATCAAGATGGGGTT 918
 QY 133 ATCCATGTTTATCGGATAAAGATTGTACGGAGAGCAATTTATCTGTGGCTGATGCTACA 192
 Db 919 GTCCACGCTGTTGCAATTAAGATGCAAAAGAACATCTTCCCGGTAGCTGATGCCACA 978
 QY 193 ACCTTCTTCACTGATGCTATATATTTCTCGAGTACGGCTGACGAGAACACAAAGAACT 252
 Db 979 GCGTTTTTCACTGATGCTACGTACGTACGTACGTACGTACGTACGTACGTACGTACGT 1038
 QY 253 GTCTGCCATTAATCGTTTAAATCTTTTGTAGCATAGTTTAAATCTCATCTCATGTTAAAT 312
 Db 1039 TTGTGCCACCGTGCATGCTCTAGAACAGAAATTTATCTCCATTTGATGCTTAAT 1098
 QY 313 GCGGATAGGAATTTCTTGCCAGAGACTGCCCCACATCGTGATTTTTTACAATGTGTCAGG 372
 Db 1099 GCGGATAAAGAAATTTCTTGCTCAAAAAGTGCACCATCGTGATTTTTTATAACGTAGG 1158
 QY 373 AAGTTGACACTCATGTTTCATTCATTCAGCATGCAATGAATCAAAAACATCTGTTGAGGTT 432
 Db 1159 AAGTTCGACACTCATGTTGTCATTCATTCAGCTTGCATGAACAGAACACCTTTTAAAGGTTT 1218

```
QY 433 ATAAATCCAACTAAGAAAAAGAACTGATGAGGTGGTCAATTTTCAGAGATGGTACTTAT 492
DB |||
QY 1219 ATTAAGTCAAGCTCGGAAAGAACCGATGAGGTGTATATTTCCGAGATGGAAATAT 1278
DB |||
QY 493 ATGACTTTTAAAGAGGTTTTTTCAGAGCTTTGGACTTAACTGGGTATGATCTGAATGTTGAT 552
DB |||
QY 1279 TTGACCTTGAGAGAAAGTTTTTTCAGAGCTTTGGATCTGACTGGATATGACCTGAAAGCTGCAC 1338
DB |||
QY 553 TTGCTAGATGTCATCGACAGCAAAAGCAATTTTCATCGTTTTTTCAGCAAAATTTCAATCTAA 612
DB |||
QY 1339 CTTTGGATGTTTCATCGACAGCAAAAGTACCTTTTCATCGTTTTTTCATGATGTTTCAACCTTAA 1398
DB |||
QY 613 TACAATCCATGTGGCCAAAGTAGGCTCAGAGAAATTTTCTCACAAGATAATCTTATTT 672
DB |||
QY 1399 TATACCTTTGGTCAAGTAGGCTTAGGAGATTTTCTTAAACAGATATCTCAATC 1458
DB |||
QY 673 CAAGGCCGTTTTCTCTGAGTTGACAAAGCAAGTTTTTCTCTGACCTTTCTCTGACAAA 732
DB |||
QY 1459 CAAGGTCGATTTCTTGGTGAGATAACAAAGCAAGTCTTCTCTGACCTTGAAGCTAGTAAA 1518
DB |||
QY 733 TATCAGATGCCAGATATAGGATTTCAATCTACGGAAGAAACAGAGTGAATGGGACCAA 792
DB |||
QY 1519 TATCAGATGGCTGAATACAGAAATATCTATATATGCGAGAAAAATGAGCGAGTGGGACCAA 1578
DB |||
QY 793 CTGCAAGTTGGATAGTGAAACAAATGAATTCGACAGTGGAAATCTTCTCTGGCTGGTTTCA 852
DB |||
QY 1579 CTCGCTAGTTGGATTTGGAACATGATCTATACAGTGAAGATTTGTTCTGGTTTAAATTCAG 1638
DB |||
QY 853 ATTCCAAGCTTATATATATGTGACAAAGAAATGGGTATCGTTACATCATTTCCAAAATCTT 912
DB |||
QY 1639 CTCCTCAAGCTTGTACAAATTTTACAAGGACATGGGTATTTGTGACATCGTTCCAGAAATATC 1698
DB |||
QY 913 CTTGACAAATTTTCTCTCTCTCTTTTGGATTTACTATTTGATCCAGCTTTCACCCACAG 972
DB |||
QY 1699 CTGGAACAATATATTTCTCTCTTGTGTAAGCAGCGTAGATCTCTGATTTCCCATCTCTAG 1758
DB |||
QY 973 CTCATGTTCTTCTGTAAGCAGGTTGTAGGTTGGACCTGTTGATGATCAATGAATGAACCA 1032
DB |||
QY 1759 CTCATGTTTTTTTGAAGCAGGTTGTGATTTGATTTGATTTGATGATGAAGCAACCT 1818
DB |||
QY 1033 GAAAGCGTCCAAACAAAGACATGCCCAACCTGAAACAGTGAACCAATGTGTTCAACCT 1092
DB |||
QY 1819 GAAAGCGTCCCAACAAACACATGCCCACTCCAGCTCAATGGACTAAGCAATTCATCTCT 1878
DB |||
QY 1093 GCATTTTTCATATATGCTGCTACTGCTATGCTTAATCTTATCCCTTAACAGCTGGCT 1152
DB |||
QY 1879 GCATTTTTCATATATGCTGCTACTGCTATGCTTAATCTTATGCTTAATTAAGCTTGA 1938
DB |||
QY 1153 GAGTCAAAAGGGAATGACCACTATCAAAATTCGGTCCACATGCTGGAGAGGCTGGAGATGTT 1212
DB |||
QY 1939 GAGTCAAAAGGGAATGACCACTATCAAAATTCGGTCCACATGCTGGAGAGGCTGGAGATGTT 1998
DB |||
QY 1213 GATCACTTGGCAGCAGATTTCTTCTCTGCAACATATCAATGAAATTAATCTTAAGG 1272
DB |||
QY 1999 GACCACTTGGCTGCTACGTTTCTTAACATGCCATAGCATCGCATGCAATCAATCTGCGA 2058
DB |||
QY 1273 AGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1332
DB |||
QY 2059 AGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2118
DB |||
QY 1333 TTGAGCAACAACTCTCTTATTTCTTGAATCATGCAACCTTTTCCAAAGCTTTCTTCCAA 1392
DB |||
QY 2119 CTGAGCAACAACTCTTGTGTTCTGATTAACCCAGGACCCGTTCTGCTGTTTCTTCTTA 2178
DB |||
QY 1393 CGAGGCTGAATGCTCTCATTTCTACGATGACCCCTTTGCAAAATTCACCTGACAAAGAA 1452
DB |||
QY 2179 AGAGGCTCTCAATGTTTCTCTGCTCTAATGATGACCCCTTTCAGATTTCACTTAACTAAGAA 2238
DB |||
QY 1453 CCATTGGTGAAGATACAGCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1512
DB |||
QY 2239 CCTCTCGTGAAGATATAGCATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2298
DB |||
```

```
QY 1513 TCGCAATTTGGAGGAACTCTGTTTACCAATCTGGGTTTTTTCACATGCTCTCAAGCGCAC 1572
DB |||
QY 2259 TGGGAGATAGCTCGTAACTCAGTGTACAGTCAGGCTTTCTCACAGCGCTGAAGTCGCAC 2358
DB |||
QY 1573 TGGATTGTGAAGAACTACTTCAAAAGAGGAGCTGCTGGAATGATATTTTCAAGAACCAAT 1632
DB |||
QY 2359 TGGATTGGAAGAAATTTACTACAAAAGAGGAGCTGATGGAACGACATTTCAAAAACAAAC 2418
DB |||
QY 1633 GTACCGCATCAGGTTTCAATTTAGAGAGATGATCTGAGAGAAATGAATTAAGAACTAGTG 1692
DB |||
QY 2419 GTCCACACATAAGGTTGGAGTTCGTGACAGATCTGGAAGAGGAGATGCAACAGTTT 2478
DB |||
QY 1693 TACTCTGACAATG 1705
DB |||
QY 2479 TATCTGGGCAAGG 2491
DB |||

RESULT 7
AY056301
LOCUS
DEFINITION
Arabidopsis thaliana putative AMP deaminase (At2g38280) mRNA,
complete cds.
ACCESSION
AY056301
VERSION
AY056301.1
GI:15810524
KEYWORDS
FLI CDNA.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicaceae; Arabidopsis.
REFERENCE
1 (bases 1 to 2847)
Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K.,
Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L.,
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,
Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T.,
Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K.,
Davis, R.W., Ecker, J.R. and Theologis, A.
Arabidopsis Full Length cDNA Clones
Unpublished
2 (bases 1 to 2847)
Yamada, K., Banh, J., Banno, F., Dale, J.M., Goldsmith, A.D., Lee, J.M.,
Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Yamamura, Y.,
Yu, G., Yu, S., Bowser, L., Carninci, P., Chen, H., Cheuk, R.,
Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G.,
Kawai, J., Kim, C., Koesema, E., Lam, B., Lin, J., Meyers, M.C.,
Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T.,
Satou, M., Seki, M., Shinn, P., Southwick, A., Tracy, S.E.,
Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
Direct Submission
Submitted (12-SEP-2001) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
Riken Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J.,
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
Hayashizaki, Y. and Shinozaki, K.

TITLE
Arabidopsis Full Length cDNA Clones
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 2847)
Yamada, K., Banh, J., Banno, F., Dale, J.M., Goldsmith, A.D., Lee, J.M.,
Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Yamamura, Y.,
Yu, G., Yu, S., Bowser, L., Carninci, P., Chen, H., Cheuk, R.,
Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G.,
Kawai, J., Kim, C., Koesema, E., Lam, B., Lin, J., Meyers, M.C.,
Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T.,
Satou, M., Seki, M., Shinn, P., Southwick, A., Tracy, S.E.,
Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
Direct Submission
Submitted (12-SEP-2001) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
Riken Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J.,
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
Hayashizaki, Y. and Shinozaki, K.

TITLE
The Salk, Stanford, PGEC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Yamada, K., Banh, J.,
Banno, F., Dale, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S.,
Quach, H.L., Tang, C.C., Toriumi, M., Yamamura, Y., Yu, G., Yu, S.,
Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C.,
Koesema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M.,
Palm, C.J., Shinn, P., Southwick, A., Tracy, S.E., Davis, R.W.,
Ecker, J.R. and Theologis, A.

TITLE
Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to
this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC)
contributed equally to this work as PIs.
Annotation is based on the January 2002 version of the Arabidopsis
genome submitted to GenBank.
```

Query Match	54.5%	Score 1041.8	DB 8	Length 2847
Best Local Similarity	76.0%	Pred. No. 9.6e-207		
Matches 1286	Conservative 0	Mismatches 407	Indels 0	Gaps 0
Qy	13	GTTCCTCCATGGGAAGAGGAGTGCATAAATGACCCCTGTACTCTCCAAACCTTAACCCCAAC	72	
Db	882	GTTCACATGGGAAAAAGAAAGTCATATCTGATCCTAGTAGTCTCCAAAGCCTAATACAGAG	941	
Qy	73	CCGTTCACTTATGTGCCTGAAACCAAAGTCAGAGCATGTTTTCCAAACTGTGTGATGGCGTT	132	
Db	942	CCATTTGCACACTATCTCTCAGGGAAAAATCTGATCATTTGTTTGAGATGCAAGATGGGGTT	1001	
Qy	133	ATCCATGTTTTATCGGATAAAGATTGTACGGAGAGCATTTATCTGTGGCTGTGATGCTACA	192	
Db	1002	GTCCACGTGTTTGCAATAAAGATGCAAAAGAGATCTTCTCCGGTAGCTGATGCCACA	1061	
Qy	193	ACCTTCTCACTGACTTGCATTATATTTCTCGAGTAAAGCGTGACGGGAACACAAAGAACT	252	
Db	1062	CGCTTTTCTCACTGACTTGTGACTACGTACTCAAAGTCACTAGCTGCAGGAAACATCCGGACT	1121	
Qy	253	GTCTGCCATAATCGGTTAAATCTCTTTGAGCATAAAGTTTAAATTCATCTGATGTTAAAT	312	
Db	1122	TTTGCCACCAGCTCGACTAGTGCTCCTTAGAACAGAAATTTAATCTCCAATTGATGCTTAAT	1181	
Qy	313	CGCGATAGGGAATTTCTTTGGCCAGAAAGCTGCCCCACATCGTGATTTTTTACAATGTCCAGG	372	
Db	1182	CGCGATAAGAAATTTCTTTGCTCAAAAAGTGCCACCATCGTGATTTTTTATACGTTAGG	1241	
Qy	373	AAGTTGACACTCAATGTTTCATTCATTCTACGATGCATGAATCAAAAACATCTGTTGAGGTTT	432	
Db	1242	AAAGTCGCACTCATGTGCTATCATTTCACTTGCATTCGATTCGATGAACACAGAAACCTTTAAGGTTT	1301	

Db 2076 GACCACTTGGCTGCTAGCTTTCTAACATGCCATAGCATCGCATGGAATCAATTCGCGA 2135
Qy 1273 AAGTCTCTGCTGCTCAGTACTTGTACTATCTTGTGTCAGATGGTCTGGCGATGTCGCCA 1332
Db 2136 AAGTCTCTGCTGCTCAGTACTTGTACTACCTCGCCAGATGGTCTGGCCATGTCACCA 2195
Qy 1333 TTGAGCAACAACCTCTTATTTCTTGACTATCATCGCAACCCCTTTCCAAAGCTTCTTCCAA 1392
Db 2196 CTGAGCAACAACCTCTTGTGTTCTAGATTACCAACCGGAACCCGTTCTCTGTTTCTTTA 2255
Qy 1393 CGAGGTCTGAATGTCTCATATCTACGGATGACCCCTTTGCAAAATCCACCTGACAAAGAA 1452
Db 2256 AGAGGTCTCAAGTCTCTGCTACTGATGATGACCCCTTCAGATTCACCTTAATAAGAA 2315
Qy 1453 CCATTGGTGAAGAAATACAGCATGTCTTCTGCTGCTGGAAGCTCAGTTCTTTGTGATTTA 1512
Db 2316 CCTCTCGTGAAGAGTATAGCATAGCTGCATCAGTTTGAAGCTGAGTGGCTGACCTG 2375
Qy 1513 TGGCAATTTGCGAGGAACCTCTGTTTACCAATCTGGGTTTTCATAGCTCTCAAGGGCAC 1572
Db 2376 TGGAGATAGCTGCTGAATCAGTGTACCATGAGTTCCTCACGCGCTGAAGTCGCAC 2435
Qy 1573 TGGATTGGTAAGAACTACTTTCAAAAGAGGACCTCTGGAATGATATTCACAGAACCAAT 1632
Db 2436 TGGATTGGAAAGATTACTACAAAGAGGACCTGATGGAACGACATTCACAAACAAAC 2495
Qy 1633 GTACCGCATCAGGGTTCATATTTAGAGAGATGATCTGGAGAAATGAAATGAAATAGTGTG 1692
Db 2496 GTGCCACACATAAGGCTGGAGTTCGTGACACGATCTGGAAGAGGAGATGCAACAGTT 2555
Qy 1693 TACTCTGACAAATG 1705
Db 2556 TATCTGGGCAAGG 2568

RESULT 9
AX507020
LOCUS AX507020 1803 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 1715 from Patent WO0216655.
ACCESSION AX507020
VERSION AX507020.1 GI:23388257
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Harper, J.F., Kreps, J., Wang, X. and Zhu, T.
TITLE Stress-regulated genes of plants, transgenic plants containing same, and methods of use
JOURNAL Patent: WO 0216655-A 1715 28-FEB-2002;
The Scripps Research Institute (US); Syngenta Participations AG (CH)
FEATURES
source
1.1803
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
BASE COUNT 524 a 384 c 378 g 517 t
ORIGIN
Query Match 53.8%; Score 1028; DB 6; Length 1803;
Best Local Similarity 76.2%; Pred. No. 7.3e-204;
Matches 1265; Conservative 0; Mismatches 395; Indels 0; Gaps 0;
Qy 13 GTTGCTCCATGGGAGAGGATCAATAATGACCCCTGCTACTCCAAAACCTTAACCCCAAC 72
Db 100 GTTGCAACATGGGAAAAGAGTATATCTGATCTCTAGTACTCCAAAGCCTTAATACAGAG 159
Qy 73 CGGTTCACTTATGTGCTTGAACCAAGTCAGAGCATGTTTCCAACTGTTGATGCGGTT 132

Db 160 CCAATTTGCACACTATCTCTCAGGAAAATCTGATCAATGTTTGTGAGATCGCAAGATGGGGTT 219
Qy 133 ATCCATGTTTATGGGATAAAGATTTGTACGGAGAGCAATTTATCTCTGTGGCTGATGCTACA 192
Db 220 GTCCACGTTGTTGCAAAATAAAGATGCAAAAAGAGATCTCTTCCCGGTAGCTGATGCCACA 279
Qy 193 ACCTTCTTCACTGACTTGCATTTATTTCTCCGAGTAACGGCTGCAGGGAACACAGAAGACT 252
Db 280 CGGTTTTTCACTGACTTGCATCAGTACTCAAGTCTCATAGCTGCAGGAAACATCCGGACT 339
Qy 253 GTCTGCCATAAATCGGTTAAATCTTCTTGAGCATAAAGTTTAAATCCATCTCATGTTTAAAT 312
Db 340 TTGTGCCACCGTGCAGTAGTGTCTCTAGAAACAGAAATTTAATCTCCATTTGATGCTTAAT 399
Qy 313 CGGATAGGGAATTTCTTGGCCGAAGACTGCCCAATCGTGATTTTAAATCAATGTGACGG 372
Db 400 GCGGATAAAGAAATTTCTGCTCAAAAAGTGCACCAATCGTGATTTTATAACGTTAGG 459
Qy 373 AAGTTGACACTCATGTTTCATCATTTCCAGCATGATGATCAAAAACATCTCTGTGAGGTTT 432
Db 460 AAGTTCGACACTCATGTTGCATCATTTCACTTTCATGAAACAGAAACACCTTTTAAGGTTT 519
Qy 433 ATAAATCCAAACTAAAGAAAAGAACCTGATGAGGTGGTCAATTTTTCAGAGATGGTACTTAT 492
Db 520 ATTAAGTCAAGCTCCGGAAGAACCCGATGAGGTTGTAATATTCGAGATGGAAACATAT 579
Qy 493 ATGACTTTAAAGAGGTTTGTGAGAGCTTGACCTTAACCTGGGTATGATCTGAATGTTGAT 552
Db 580 TTGACCTTTGAGAGAAGTTTGTGAGAGCTGGATCTGACTGGATATGACCTGAACGTCGAC 639
Qy 553 TTGCTAGATGTCCATGACAGACAAAAGACATTTTCATCGTTTTTGACAAATTCATCTAAAA 612
Db 640 CTTTTGGATGTTTCATGACAGACAAAAGTACCTTTTCATCGTTTTGTAGTTCACACCTAAG 699
Qy 613 TACAATCCATGTGGCCAAAAGTAGGCTCAGAGAAATTTTCTCAACAGATGATTAATCTTAT 672
Db 700 TATAACCTTGTGTCAAGTAGGCTTAGGAGATTTTCTTAAACAGGATTAATCTCATC 759
Qy 673 CAAGCCGTTTCTTGTGCTGAGTTGACAAAGCAAGTTTTTCTCTGACCTTTCTGTAGCAAAA 732
Db 760 CAAGGTGATTTCTTGTGAGATAAACAAGCAAGTCTTCTCTGACCTTTGAAGCTAGTAAA 819
Qy 733 TATCAGATGGCAGAAATATAGGATTTCAATCTACGGAAGGAAAACAGATGAAATGGACCAA 792
Db 820 TATCAGATGGCTGAATACAGAAATATCTATATATGCGAGAAAAATGACGAGTGGGACCAA 879
Qy 793 CTTCAGAGTTGAGTAGTGAACAATGACAGTGCAGAAATGTTGTTGCTGGCTGGTTCAG 852
Db 880 CTCGCTAGTTGGATTGTGAACAATGATCTATACAGTGAAGATTTGTCTGGTTTAAATCAG 939
Qy 853 ATTCACGCTTATATAATGTGTACAAGAAATGGGTATCGTTACATCATTTCCAAAATCTT 912
Db 940 CTCCACGCTGTGACAAACATTTACAGGACATGGGTATTTGATGATCGTTCCAGAAATC 999
Qy 913 CTTGACAAACATTTTGGTTCTCTTTTGTGAGTTACTATTTGATTCAGCTTTCACCCACAG 972
Db 1000 CTGGACAAATATATATTCATTTCTCTGTTTGAAGCCACGGTAGATCTCTGATTCCTCATCCTCAG 1059
Qy 973 CTCCATGCTCTTCGAGGAGGTTGTAGGTTGGACCTGGTTGATGATGAAGTAAGAACCA 1032
Db 1060 CTCCATGTTTGTGAGGAGGTTGTGGATTTGATTTGGTTGATGATGAAGGAAACCT 1119
Qy 1033 GAAAGGGGTCCAAACAAAGACACATGCCACACCTGAAACAGTGGACCAATGTGTTCAACCCCT 1092
Db 1120 GAAAGAGCTCCCAACAAACACATGCCCACTCCAGCTCAATGGATTAACGCATTCATCTCT 1179
Qy 1093 GCATTTTCATATATGCGTACTACTGCTATGCTTAATTTTCAACCTTAAACAAAGCTGCT 1152
Db 1180 GCATTTTCATATATGCTACTATTTGTTATGCTAACCTCTACTGTGTTTAAATAGCTTCGA 1239
Qy 1153 GAGTCAAGGGAATGACCACTATCAAAATTCGTTCCACATGCTGGAGAGGCTGGAGATGTT 1212
Db 1240 GAGTCAAGGGAATGACTACTATCACTATCACTGACCAACCACTTCTGGAGAGGCTGGTACATT 1299

```

QY 1213 GATCACTTGGCAGCGACATTTCTTCTCTGTCAACATATACATCGGAATTAATCTAAGG 1272
Db 1300 GACCCTTGGCTGCTACGTTTCTAACATGCCATAGCATCGCATGGAATCAATCTGCGA 1359
QY 1273 AAGTCTCTCTGTGCTTCAGTACTTGTACTCTTGTGTCAGATTGGTCTGGCGATGTCGCCA 1332
Db 1360 AAGTCTCTCTGTGCTTCAGTACTTGTACTCTTGTGTCAGATTGGTCTGGCGATGTCGCCA 1419
QY 1333 TTGAGCAACAATCTCTTATTTCTTGACTATCATCTGCAACCCCTTTTCCAAAGCTTTCTTCCA 1392
Db 1420 CTGAGCAACAATCTCTTGTGTTCTAGATTACCAACCGGAACCCGTTTCTGTGTTTCTTCTTA 1479
QY 1393 CGAGGTCGAATCTCATATCTACGATGACCTCTTGCATATTCGCAATTCACCTGACAAAGAA 1452
Db 1480 AGAGGTCCTCAATGTTTCTCTGCTACTGATGACCCCTTCAGATTCACTTAATAAGAA 1539
QY 1453 CCATTGGTGGGAAGATACAGCATCTGCTTGTGCTGGAAGCTCAGTTCTTGTGATTTA 1512
Db 1540 CCTCTCTGGAAGATATAGCATAGCTGCATCAGTTTGGAGCTGAGTGGTGTGACCTG 1599
QY 1513 TGGGAATTTGGCAGGAATCTCTGTTTACCAATCTGGGTTTTCATGCTCTCAAGGCGAC 1572
Db 1600 TGGCAGATAGCTCGTAACCTCAGTGTACCACTCAGTGTGTTCTCACACGCCCTGAAGTGGCAC 1659
QY 1573 TGGATTGGTAAGAACTACTTCAAAAGAGGACCTGCTGGAATGATATTCACAGAACCAAT 1632
Db 1660 TGGATTGGAAAGATTACTACAAAGAGGACCTGATGGAACGACATTCACAAAACAAAC 1719
QY 1633 GTACCCACATCAGGGTTCATATTTAGAGAGATGATCTGGA 1672
Db 1720 GTGCCACACATAGGGTGGAGTTCCTGTGACACGGTATGGA 1759

```

RESULT 10

```

AX081198
LOCUS 2482 bp DNA linear PAT 27-FEB-2001
DEFINITION Sequence 17 from Patent WO0109305.
ACCESSION AX081198
VERSION AX081198.1 GI:13170074
KEYWORDS Glycine max (soybean)
SOURCE Glycine max
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1
REFERENCE
AUTHORS Caspar, T., Falco, S. C., Sakai, H., Weng, Z. and Hu, X.
TITILE Purine metabolism genes in plants
JOURNAL Patent: WO 0109305-A 17 08-FEB-2001;
E.I. DU PONT DE NEUMOURS AND COMPANY (US) ; PIONEER HI-BRED
INTERNATIONAL, INC. (US)
FEATURES
Location/Qualifiers
source 1..2482
/organism="Glycine max"
/mol_type="genomic DNA"
/db_xref="taxon:3847"
BASE COUNT 688 a 539 c 569 g 686 t
ORIGIN

```

Query Match 50.4%; Score 962.4; DB 6; Length 2482;

Best Local Similarity 78.7%; Pred. No. 3.3e-190;

Matches 1149; Conservative 0; Mismatches 311; Indels 0; Gaps 0;

```

QY 13 GTTGCTCCATGGGAGGAGGATCAATAATGACCCCTGTACTCCAAACCTTAACCCCAAC 72
Db 1023 GTTGCTCCGTTGGATTAAGATTATATCCGACCCGAGCACACCAAGCTTAACCCAGAT 1082

```

```

QY 73 CGGTTCATCTATGTGCTGAAACCAAGTCAGAGATGTTTTTCCAACTGTGTGATGGGTT 132
Db 1083 CCATTTTATACATTTCTGAAGGAATTTCTGATCATTTATTTTGAATGCAAGATGGGTT 1142

```

```

QY 133 ATCCATGTTTATGCGGATAAAGATGTACGGAGAGCATTTATTCCTGTGGCTGATGCTACA 192
Db 1143 ATTCTGTATATACAGATAGAGATCCAAAGAGAGCTTTTTCTGTAGCCGATCAACT 1202
QY 193 ACCTTCTTCACTGACTTGCATATATTTCTCGAGTAACGGCTGCAGGAAACACAAGAACT 252
Db 1203 ACATTTTTCACTGATCTTCACTACTTCCAGTATAGCAGCAGGAGGAATATAAGAACT 1262
QY 253 GTCTGCCAATACCGGTTAAATCTTCTTGAGCATAGTTTAAATCCATCTGATGTTAAAT 312
Db 1263 TTATGCCATCATAGGCTCAATCTTCTAGAACAAAATCAATCTTCAITGATGCTAAAT 1322
QY 313 GCGGATAGGGAATTTCTTCCAGAGAGACTGCCACATCGTGATTTTATCAATGCTCAGG 372
Db 1323 GCGGATAGAGAAATTTCTTCTCAGNAGAGTGTCCACATCGAGACTTCTATAATGTTAGA 1382
QY 373 AAGTTGACATCTATGTTTCATCTCAGCATGCAATCAAAAACATCTCTGTGAGGTTTC 432
Db 1383 AAGTTGATCTCATGTCCACACTCAGCATGCAATCAAAAACATCTCTTAAAGGTTTC 1442
QY 433 ATAAATCCAAACTAAGAAAGAACCTGATGAGGTGCTCATTTTCAGAGATGGTACTTAT 492
Db 1443 ATAAAGTCAAAAGCTGAGAAAGAGCCTGATGAGGTGTTAAATATTTTCGAGATGGGACATAT 1502
QY 493 ATGACTTTAAAGAGGTTTTTGGAGAGCTTGGACTTAACTGGGTATGATCTGAAATGTTGAT 552
Db 1503 CTAAAGTTGGAGAGGTTTTCAAGAGTTTAGATTGTCTGGATATGACCTCAATGTTGAC 1562
QY 553 TTGCTAGATGTCATGCAGACAAAAGCACATTTTCATCGTTTTTGGACAAATTCATCTAAA 612
Db 1563 CTTTTGGAGCTTCAACGACAGCAAGAGTACTTTTTCATCGTTTGAAGTTCAATCTTAAA 1622
QY 613 TACAATCCATGTGGCCAAAGTAGGCTCAGAGAAATTTTCTCAAACAGATATCTTAT 672
Db 1623 TACAATCTCTGCGSTCAAAGTAGGCTCAGGAGATATTTCTTAAACAGGATATCTCAT 1682
QY 673 CAAGCCGTTTTCTTGTGAGTTGACAAAGCAAGTTTTCTCTGACCTTCTGCTAGCAAA 732
Db 1683 CAAGTCTGTTTTCTTGTGAGTTTAACTAAGCAAGTTTTTTCAGATCTTCTGCTAGTAAA 1742
QY 733 TATCAGATGGCAAAATATAGGATTTCAATCTACGGAAGAAACAGAGTGAATGGGACCAA 792
Db 1743 TATCAGATGGCTGAAATATAGAAATCAATATATGTAGGAAGCAAGTGAAGTGGGACCAA 1802
QY 793 CTTGCAAGTTGGATAGTGAACAAATGCAACAGTGGAAATGTTGCTGCTGGCTGCTGAG 852
Db 1803 CTAGCCAGTTGGATAGTGAATATGATTTGTACAGGAGATGTCTGTTGTTGATTCAG 1862
QY 853 ATTCACGCTTATATAATGTGTACAGGAAATGGGTATCGTTTACATCATTTCCAAAATCTT 912
Db 1863 CTTCCACGGTTGTACAATGTGTACAAGAAATGGGAATGTGACATCATTTCCAGAACATG 1922
QY 913 CTTGCAACATTTCTGCTCTCTTTTGGGTTACTATTGTATCCAGTTCACACCCACAG 972
Db 1923 CTCGACAATATTTTCATTTCCACTTTTGGAGTCACTGTCAACCCAGATTCACATCTCAG 1982
QY 973 CTTCAATGCTTCTCAAGCAGGTTGTAGGTTGGACCTGGTTGATGATGAAGTAACCA 1032
Db 1983 CTTGCAATGTTTCTCTGAAACAGGTTGTGGGTTTGGATTTGGTTGATGATGAAGAACCT 2042
QY 1033 GAAAGCGCTCCAAAGCACATGCCACACCTGAAACAGTGGACCAATGTGTTTCAACCT 1092
Db 2043 GAAAGCGCCCAAAACACATGCTACACCTGAGCAATGACATGATGTTTCAATCCG 2102
QY 1093 GCATTTTCAATATTATGCGTACTGCTATGCTAACTTATTCACCTTAAACAGTGGCT 1152
Db 2103 GCATTTTTCATATGCTTACTTGTATGCAAAATCTTTACACCTTAAACAGTTCGA 2162
QY 1153 GAGTCAAGGGAATGACCACTATCAATTTCCGTCACATGCTGGAGAGGCTCGAGATGTT 1212
Db 2163 GAATCAAGGGAATGACAAATCAATTTCCGTCACATTTCTGGAGAGGCTGGTGATAT 2222
QY 1213 GATCACTTGGCAGCGACATTTCTTCTGTCTCAACATATCATATGGAATTAATCTAAGG 1272

```



```

RESULT 12
AX357098
LOCUS AX357098 914 bp DNA linear PAT 13-FEB-2002
DEFINITION Sequence 17 from Patent WO0206319.
ACCESSION AX357098
VERSION AX357098.1 GI:186744279
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1
Schulz, A., Streber, W., Hanke, C., Schmidt, F. and Schubel, A.
Nucleic acid molecule which codes for a plant amp deaminase
Patent: WO 0206319-A 17 24-JAN-2002;
Aventis CropScience GmbH (DE)
Location/Qualifiers
1. .914
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
<2..880
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD23277.1"
/db_xref="GI:18674280"
/translation="PRLYNIYKDMGIVTFSQNILNDIFILPEATVDPDSHPOLHVFL
KQVGFVLDVDES KPERRTKMPPTAOWTDFNPAFYVYVYCYANLYVNLKJRESK
GWTITLRLPHSGEAGIDHLAATFLTCHSIAGHINLRKSPVLQYLYLAQIGLWMSPL
SNNSFLDVRHPFPVFFLRGLNLSLTDPLQIHLEKPLVEEYSIAASVWKLACD
LCEIARNSVYQSGFSLKSHWIKGRDYKRGPDGNDIHKTNVPHIRVFRFTWKEEM
QQVILKAVISDEVP"
BASE COUNT 250 a 215 c 196 g 253 t
ORIGIN
Query Match 27.8%; Score 531.6; DB 6; Length 914;
Best Local Similarity 76.8%; Pred. No. 1.7e-100;
Matches 651; Conservative 0; Mismatches 199; Indels 0; Gaps 0;
QY 856 CCACGCTTATATATGTGTAAGGAAATGGGTATCGTTACATCATTCACAAATCTTCTT 915
Db 2 CCACGCTTGACAACTTTACAGGACATGGGTATTGTGACATCGTTCCAGATAATCTCTG 61
QY 916 GACAACTTTCTCTCTCTTTTGGAGTTACTATTGATCCAGCTTCACACCCACAGCTC 975
Db 62 GACAATATATATCTCTCTCTTTTGAAGCACCGGTAGATCCTGATTCCTCATCTCAGCTC 121
QY 976 CATGTCTTCTGTAAGCAGGTTGTAGGTTGGACCTGTTGATGATGATGATGATGATGATG 1035
Db 122 CATGTTTTTTTGAAGCAGGTTGTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 181
QY 1036 AGCGCTCCACAAAGCACATGCCACACCTGAAACAGTGACCAATGTGTTCAACCTGCA 1095
Db 182 AGAGCTCCCAACAAACACATGCCACTCCAGCTCAATGAGTCAAGCATTCATCTGCA 241
QY 1096 TTTTTCATATATCGGTACTACTGCTATGCTATCTTATTCACCTTAACAGCTGCGTGAG 1155
Db 242 TTTTTCGTTATATGCTACTATTGTTTATGCTTAACCTCTATGTTGTTAAATAGCTTCGAG 301
QY 1156 TCAAGGGAATGACCACTATCAATTCGCTCCACATGCTGAGAGGCTGGAGATGTTGAT 1215
Db 302 TCAAGGGCATGACTACTATCAGCTACGACCACTTCCTGAGAGGCTGGTGCATTTGAC 361
QY 1216 CACTTGCCAGCAGATTTCTTCTGTGACAAACATATACATGGAATTAATCTAAGGAAG 1275
Db 362 CACTTGCTGCTAGCTTTCTTAACATGCCATAGCATCGCATGGAATCAATCTGCGAAG 421
QY 1276 TCTCCTGTGCTTCAAGTCTTACTACTTCTTGTGTCAGATTTGCTGGGATGTCCTCATTTG 1335
Db 422 TCTCCTGTGCTTCAAGTCTTACTACTCTCTCTGCGCCAGATTTGCTGGCCATGTCACCACTG 481

```

```

QY 1336 AGCAACAACTCCTTATTTCTTGACTATCATCGCAACCCCTTTTCCAAACGTTCTTCCAACGA 1395
Db 482 AGCAACAACTCCTTATTTCTTGACTATCATCGCAACCCGTTTCTCTGTCTTTTCTTAAGA 541
QY 1396 GGTCTGAATGCTCATTTATCTACGGATGACCTTTGGAAATTCACCTGACAAAGAACA 1455
Db 542 GGTCTCAATGTTCTCTCTACTGATGACCCCTTTCAGATTTCACTTGAAGAAACCT 601
QY 1456 TTGTTGGAAGAATACAGCAATTCGCTTGGTGGAAAGCTCAGTTCTTGTGATTATGTC 1515
Db 602 CTCGTGGGAAGATATAGCATAGCTGATCTGTTGGAAAGCTGAGTGGTGTGACCTGTGC 661
QY 1516 GAAATTGCGAGAACTCTGTTTACCAATCTCGGTTTTCACATGCTCTCAAGGCGCACTGG 1575
Db 662 GAGATAGCTCGTAACTCAGTGTACAGTCAGGTTTCTCACACGCCCTGAAGTCGCACTGG 721
QY 1576 ATTGTTAGAACTACTTCAAAAGAGGACCTGCTGGAATGATATTTCACAGAACCAATGTA 1635
Db 722 ATTGGAAGAAGATTACTACAAAAGAGGACCTGATGGAACGACATTCACAAAACAAACGTTG 781
QY 1636 CGGCACATCAGGTTCAATTTTAGAGAGATGATCTGGAGAAATGAAATGAAACTAGTGTAC 1695
Db 782 CCACACATAAGGTGGAGTTCCGTGACACAGATCTGGAAGAGGAGATGCAACAGGTTTAT 841
QY 1696 TCTGACAATG 1705
Db 842 CTGGGCAAGG 851

SPBC106 42544 bp DNA linear PLN 08-DEC-1999
S. pombe chromosome II cosmid c106.
AL110295.2 GI:6562182
60s ribosomal protein 125-a; adal; Adenosine/AMPdeaminase; amp
deaminase; camp-dependent protein kinase; cct4; chaperonin;
chaperonin subunit cct4; cut4; exocyst complex; homoserine;
o-acetyltransferase; idl1; isopentenyl-diphosphate delta-isomerase;
mph1; pkai; PROS28 family; proteasome component; protein kinase
homolog; RNA binding protein; rpi25a; tpk.
Schizosaccharomyces pombe (fission yeast)
Schizosaccharomycetes;
Schizosaccharomycetaceae; Schizosaccharomycetaceae;
Schizosaccharomycetes.
1 (bases 1 to 42544)
Lyne, M., Rajandream, M.A., Barrell, B.G., Davis, P. and Churcher, C.M.
Submitted (26-AUG-1999) European Schizosaccharomycetes genome
sequencing project, Sanger Centre, The Wellcome Trust Genome
Campus, Hinxton, Cambridge CB10 1SA, E-mail: barrrell@sanger.ac.uk
On Dec 12, 1999 this sequence version replaced gi:5817266.
Notes:
Details of yeast sequencing at the Sanger Centre are available on
the World Wide Web.
(URL, http://www.sanger.ac.uk/Projects/S_pombe/)
During 1995 to 1996 about 66% of S. pombe chromosome 1 was
sequenced by the Sanger Centre. The sequencing of the S. pombe
genome is now being continued with funding from The European
Commission. Fourteen European sequencing laboratories, including
the Sanger Centre, are participating in the project.
protein coding regions (CDS) have been predicted with the help of
computer analysis using the GeneFinder program in PomBase (an ACEDB
database) with additional predictions for the branch-acceptor sites
supplied by the program Splice. CAUTION: It is possible that for
any individual CDS we may have underestimated or overestimated the
number of introns/exons or we may not have chosen the correct
splice donor/acceptor sites.
CDS are numbered using the following system eg SPBC25H2.01c. SP (S.
pombe), B (chromosome 2), c25H2 (cosmid name), .01 (first CDS), c
(complementary strand).
The more significant matches with motifs in the PROSITE database

```


KKIKAGANVILIQSIILRDVNDLALHFLAKLIMVINKDIERDEVEFICKSTGCKPI
ADIESFAEDKHADHVEDTSSSEKIKVFSKVNAGKTVSILCRGANLLTLEAEARS
LHDLVCVRLVKORALJAGGSPEIARORLLEHAROLEGEALICIRAFSEALRIIP
VTLAENAGNALQVTELSRHRANGKTAGINVRKGVITNILEENVLPBLVNISAIO
LAAETTKIMIKIDITTLAR"
8768. .10246

misc_feature

/gene="SPBC106.06"
/note="Match to PF00118 cpn60_TCP1, TCP-1/cpn60 chaperonin
family Score 634.60"
complement(10863. .11366)
/gene="SPBC106.07c"
complement(10863. .11366)

Query Match 26.1%; Score 498.6; DB 8; Length 42544;
Best Local Similarity 60.1%; Pred. No. 1.1e-93; Indels 0; Gaps 0;
Matches 828; Conservative 0; Mismatches 549;

QY	205	GACTTCGATATATATCTCGAGTAACGGCTGCAGGGAACACAGAACTGCTGCCATAAT	264
DB	5584	GATTTAGAATTCCTCTTCGGCTTCATCGGATGGTCTAGTAATCAATTTTCCTTCCTG	5643
QY	265	CGGTTAAATCTCTTGAGCATATAGTTTAAATCCATCTCATGTTAAATCGGATAGGAA	324
DB	5644	AGATTGCAGTATCTAGAAGGACGCTGGAATATATATGCTGCTAAACGAGTATCAAGAG	5703
QY	325	TTTCTTGCCAGAGACTGCCACATCGTGATTTTACAATGTCAGGAAGGTTGACACT	384
DB	5704	CTGGCAGATACCAAAAAGTACCCCATCGTATTTTCAATGATAGTAAGTAGTACT	5763
QY	385	CATGTTTCATCATTCAGCATGCATGAATCAAAAAACATCTGTGTAGGTTCTATAAATCAAA	444
DB	5764	CAGTCCATCATCTCTCTGGAATCAAAAACATCTCTTGGCATTTATCAAGCTAAA	5823
QY	445	CTAAGAAAGAACCTCATGAGGTGGTCATTTTCAGAGATGGTACTTATATGACTTTAAA	504
DB	5824	CTTCGTAATGCCCTAATGAAAAGTATATCTGCGGTGATGGCAAAATCTCTGACTTTTACAA	5883
QY	505	GAGGTTTTGAGAGCTTGGAATTAACGTGGTATGATCTGAATCTGATTTGCTAGATGTC	564
DB	5884	GAAGTATTTGCTGTTGAAGTTGACTTCGTATGACCTTTAGTATAGACAGCTTGGATATG	5943
QY	565	CATGCAGACAAAAGCACATTTCACTGTTTTTGACAAATTCAAATCAAAATCAATCATGT	624
DB	5944	CATGCTCATACTGACACCTTTTCAACGATTTGACAAATTTAACTTGAATATATATCAAT	6003
QY	625	GGCCAAAGTAGGCTCAGAGAAATTTTCTCAAAACAGATATCTTTATTCAGCCGCTTTT	584
DB	6004	GGTGAATCTCGTTTGGCTACTATTTTCTGAAACCGGATAATGATATTAACGCTCGATAT	6063
QY	685	CTTGCTGAGTTGACAAAGCAAGTTTCTCTGACCTTCTGCTAGCAAAATATCAGATGGCA	744
DB	6064	CTTGCTGAACTTACAAAGAGTTGTTTACCGACCTTTAGAACTCAAAAGTACCAATGGCT	6123
QY	745	GAATATAGGATTTCAATCTACGGAAGAAACAGAGTGAATGGGACCAACTTGCAGAGTTGG	804
DB	6124	GAATATCGTATTTCTATCTATGACGGAATCGTGAAGATGGGTAAGCTTGACGCTGG	6183
QY	805	ATAGTGACAAATGAATTCAGTGGAAATGTTGCTGGCTGGTTGAGATTCACACGCTTA	864
DB	6184	ATTATTGATTAACGAACCTTTTTTCTCCCAACGTTTCTGGCTTATTAAGTACCTCGTTTG	6243
QY	865	TATAATGTGTACAAAGGAATGGGTATGTTTACATCATTTCCAAATCTCTCTTGACACAT	924
DB	6244	TATGATGTGTATGAAGATCGGTTATGTTGAGACTTTGAAGAGGTGCTGCAGAAATGTC	6303
QY	925	TTGCTTCTCTTTTGGAGTTACTATTATGATCCAGCTTTCACACCAAGCTTCCATGCTTC	984
DB	6304	TTTGAACCACTGTTTGAAGTTTACAAAGGATCTCTGACACATCTTAACTCCACGCTTT	6363
QY	985	CTGAGCAGGTTGTAGGTTGGACCTGGTTGATGATGAAGTAAACAGCAAAAGCGCTCCA	1044
DB	6364	TTGCAACGTGTATTCGGGTTTGAATCTCTGTTGACGACGAAAGTAAACCTGAACCGCTACC	6423

QY	1045	ACAAAGCACATGCCACACACCTGAAACAGTGGACCAATGTGTTCAACCTCGCATTTTCATAT	1104
DB	6424	TTTCGAAATTCCTTATCCGAAACATGGGATATTAATTTAAATTCCTCCATAGTTAT	6483
QY	1105	TATGGTACTACTGCTATGCTAACTTATTCACCCCTAAACAGCTCGGTGAGTCAAAAGGA	1164
DB	6484	TGGTTATATTATCATGTATGCAACATGACTAGCTTGATAGCTGGAGAAAATACGTGGT	6543
QY	1165	ATGACCACATCAAAATTCGGTCCACATGCTGGAGAGGCTGGAGATGTTTGATCAGTTCGCA	1224
DB	6544	TTTAAACACCTTTGCTTCCTCGTCCTCATTTGCTGAGCTGGAGATACGATCAITTTGCC	6603
QY	1225	GGCAGATTTCTCTCTGTCAACACATATCATATGGAATTAATCTAAGAAAGTCTCTCTGT	1284
DB	6604	AGTCTTTCTTATTAAAGTCACGGTATTAAACATGATTTTATTACGAAAAGTTCCTTTT	6663
QY	1285	CTTCAGTACTTGTACTATCTTTGGTCAGATTTGGTCTGGCGATGTCCTCCATTCAGCAAC	1344
DB	6664	CTTCAGTATCTTTGGTACTTTGATCAGATCCCTATCGCAATGTCTCCATTAATCAACAAT	6723
QY	1345	TCCTTATTTCTGACTATCATCGCAACCTTTTCCACGTTCTTCCACAGAGGTCTGAAT	1404
DB	6724	GCAGTCTTTCTGATATGATAGATCCATTTTGGTACTTCAACAGAGGTCTTAAT	6783
QY	1405	GTCTCATTTATCTACGATGACCTTTGCAAAATTCACCTGACAAAAGAACCAATTTGGTGAA	1464
DB	6784	GTAGTCTTTCCACTGATGATCTTTTACAATTTGCAATTTACCCGAGAGCTCTGATTGAG	6843
QY	1465	GAATACAGATTTGCTGCTTCGCTGGAGCTCAGTTCTTGTGATTTATGCAATTTGGT	1524
DB	6844	GAATATGCTGTAGCGGCACAAATTTATAAGCTCTCTGCTGCGATATGCGAATTTGGCT	6903
QY	1525	AGGAATCTGTTTACCAATCTGGGTTTTCACATGCTCTCAAGGCGCATCTGGAATTCGT	1581
DB	6904	AGAAACAGTGTTTTCAATTCGGTTTCGAACTGCTGAGCTTAAGAACGTTGGTTGGT	6960

RESULT 14

BC049119

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

BC049119 3667 bp mRNA linear ROD 22-APR-2003
Mus musculus adenosine monophosphate deaminase 2 (isoform L), mRNA
(cDNA clone MGC:61170 IMAGE:6812571), complete cds.

BC049119

BC049119.1 GI:29145072

MGC

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 3667)

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,

Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,

Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,

Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,

Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,

Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,

Carrinzi, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,

Armspith, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,

Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,

Fahney, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,

Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,

Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,

Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,

Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E.,

Schmerch, A., Schein, J.E., Jones, S.I. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

12477932

REFERENCE
AUTHORS
TITLE
JOURNAL

2 (bases 1 to 3667)
Strausberg,R.
Direct Submission
Submitted (19-MAR-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabbs@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr.
Thomas L. Casavant.
Web site: <http://genome.uiowa.edu>
Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
Bonaldio,M.F., Akabogu,I., Bair,T., Bair,J., Crouch,K., Davis,A.,
Fishler,K., Keppel,C., Kucaba,T., Lebeck,M., Melo,A., Schaefer,K.,
Scheetz,F., Smith,C., Snir,E., Tack,D., Trout,K., Walters,J.,
Casavant,T., Soares,M.B.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series:

Plate: Row: Column: 0.

FEATURES
source

Location/Qualifiers
1..3667
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="MGC:61170 IMAGE:6812571"
/tissue_type="Brain, mouse, 13.5, 14.5, 16.5, 17.5 dpc"
/clone_lib="NIH BMAP_FW0"
/lab_host="DH10B"
/note="Vector: pYX-ASC"

gene

1..3667
/gene="Ampd2"
/note="synonyms: Ampd-2, 1200014F01rik"
/db_xref="LocusID:109674"
/db_xref="MGI:88016"

CDS

401..2797
/codon_start=1
/product="Ampd2 protein"
/protein_id="AAH49119.1"
/db_xref="GI:29145073"
/db_xref="LocusID:109674"
/translation="MASEARSLGASPLQASRLPQAPCLKHPFLDLRTSMGKCKE
IAELFSLSELSRSPAYEPFESPIQLEERRQRLERISQDVKLEPDLILRAKQ
DFLKTSDSDLOLYKEQEGQDGLWEDVVLREFORVILSGEKGCVPTDLDA
AKSVVRLFIKRYMALSLQSPTRRYLQOLAEPKLETRTYEQSPDTPVSADAPVH
PPALQHPYHCEPSPAMPDGLGLRMVGVVHVTRDPDHECPVELPYPDLQEFV
ADVNLMALINGPIKSFYRLOVLSKFQMHVLLNEMKELAAOKVPHRDFYNIRK
VDTHIHASCNNQKHLRFIKRMRHLEIEIVHVEQREQTLRFVESMNLTAIDLVS
DILDVAERLTHREDKFNKPNPIGESVLRIFIKDNKISGKYFAHIIKEVMADLE
ESKYQNALRSILYGRSDNDKLARVNVHNSPVRWLQVPLRFDVYRTKGOLA
NFQEMLENIPLFEPATVPASHPELHLFLEHVDGDFSVDDESKENPHEVFNLSPLPE
AWVEDNPYPAYLYVTYANMALNHLRORGFHTFVLPHCGEAGPIHILVSAFMLA
ENISHLGLLRKAPLYLYLAQIGIAMSPLSNLSLFLSYHNPPLPEYLSGLMVL
TDDPLQFHTKEPLNEEVSIAQVWKLSCDMCLARNLSVMSGFSHKVSHWLGPNY
TKEGPEGNDIRKTPVDIRVGYRTGGLTCCOELALITQAVQSEMLETIPBEVGIVMSPG
"

BASE COUNT
ORIGIN

759 a 1051 c 1030 g 827 t

Query Match 25.5%; Score 487.8; DB 10; Length 3667;

Best Local Similarity 58.2%; Pred. No. 2,1e-91;
Matches 877; Conservative 0; Mismatches 627; Indels 3; Gaps 1;

Qy

146 CGGATAAGATTGTACGGAGAGCATTTATCTCTGGCTGATGCTACAACTTTCTTCACTG 205

Db

1173 CTGATGAGCACTGTCCGGAGGTGGAGCTTCCATACCTGACCTACAGAAATTGTAGCTG 1232

Qy 206 ACTTGCAATTATTTCTCGAGTAAGCGCTGAGGGAAACAAGAACTCTCTGCGCAATAATC 265
Db 1233 ACGTCAATGTGCTGATGCCCTGATCATCAATGGTCCCATAAAGTCATTTCTGCTACCGCC 1292
Qy 266 GGTAAATCTTCTTGAGCATAAAGTTAAATCCATCTGATGTTAAATCGCGATAGGGAAT 325
Db 1293 GGCTGAGTACTGAGCTCCAAATTCAGATGCGATGCTTTTCTCAATGAGATGAAGGAGC 1352
Qy 326 TTCTTGCCAGAGAAGCTGCCCCACATCGTGATTTTTTACAATGTCAGGAAGGTGTGACATC 385
Db 1353 TCGTGTCTCAGAGAAAGTGCACACCGGACTTCTACAATATCGTAAGGTGGACACAC 1412
Qy 386 ATGTTTCATCTCAGCATGCAATCAAAACATCTGTTGAGGTTCTATAAAATCCAAAC 445
Db 1413 ACATCCACGCTCGTCTGTCATGAACCAAGAAACATCTACTCGCTTCTCAAGCGGGCCA 1472
Qy 446 TAAGAAAAGAACCTGATGAGCTGTGCTATTTTCAGAGATGGTACTTATATGACTTTAAAG 505
Db 1473 TGAAGCGGCACCTGGAGGAGATTGTGCATGTGGAAAGCGGCGGAGCAGACGCTGAGAG 1532
Qy 506 AGGTTTTTGAGAGCTTTGAGCTTAACTGGGTATGATCTGAATGTTGATTTGCTAGATGTC 565
Db 1533 AAGTCTTCGAGAGCATGAACCTCACTGCTACGACTTAAGTGTGACACGCTGGATGTC 1592
Qy 566 ATGACAGCAAAAGACATTTTCATCGTTTGA CAATCAATCAATCAATCAATCAATCAAT 625
Db 1593 ATGCGGACAGGATACCTTTTCATCGATTGTA CAATCAATCAATCAATCAATCAATCAAT 1652
Qy 626 GCCAAGTAGCTCAGAGAAATTTTCTCAACCAAGATTAATCTTATCAAGCGCGTTTTTC 685
Db 1653 GGGAGTCTGTTCTCCGAGAGATCTTCATTAACACCGACAAACAAGATTTCTGGGAAGTACT 1712
Qy 686 TTGCTGAGTTGACAAAAGCAAGTTTTTCTCTGACCTTTCTGTAGCAAAATATCAGATGGCAG 745
Db 1713 TTGCTCAGATCATCAAGGAGGTGATGCGAGACTTGGAGGAGCAATACAGATGCGAG 1772
Qy 746 AATATAGATTTCATCTACGAAAGGAAACAAGATGAATGGGACCAACTTGCAGAGTTGGA 805
Db 1773 AGCTCCGCTGTCCATCTACGCGGCTTCGAGGGATGAGTGGGACAAGCTGCGACGCTGG 1832
Qy 806 TAGTGAAACAATGAATTCGACAGTGGAAATGTGCTGCTGCTGCTTTCAGATTCACGCTTAT 865
Db 1833 CAGTGAACCAACAAGTGCACCTCTCCCAATGTCCGCTGGCTGGTGCAGGTGCGCGCTCT 1892
Qy 866 ATAATGTGTACAAGGAAATGGGTATCGTTACATCATTTCCAAAACTCTTCTTGACAACTTT 925
Db 1893 TCGATGTGACCGCACCAAGCGGCGAGCTGGCCAACTTCCAAGAGATGCTGAGAACATCT 1952
Qy 926 TCGTTCTCTTTTGGAGTTACTATTGATCCAGCTTCAACCCACAGCTCCATGCTTTCC 985
Db 1953 TTCTGCGCTGTTTGGAGCTACTGTGCACTGCACTGCGACACCGGAGCTGCACCTCTTTC 2012
Qy 986 TGAAGCAGTTGTAGGTTGGACCTGTTGATGATGAAGTAAACCAAGAAAGGCGTCCA 1045
Db 2013 TGAAGCAGCTGGATGGTGTGATGCGTGGATGATGATGATGATGATGATGATGATGATGAT 2072
Qy 1046 CAAA---GCATGCCACACCTGCAACAGTGGACCAATGTGTTCAACCTCTCATTTTTCAT 1102
Db 2073 TCAACCTGGAGAGTCCCTCCAGAGCTTGGGTGGAGGAGGACAACTCTCTCTATGCT 2132
Qy 1103 ATTAGCGTACTACTGCTATGCTAACTTATTCACCTTAAACAGCTCGGTGAGTCAAGG 1162
Db 2133 ACTACCTGTACTACACTTGGCTAACATGGCTATGTTGAACCATCTGCGCAGGCGAGAG 2192
Qy 1163 GAATGACCATATCAAAATTCGTTCCATGCTGAGAGGCTGGAGATGTTGATCATTG 1222
Db 2193 GTTTCACACAGCTTCGCTGAGCGCGCACTGTGGGAGGCGCGGCCCATCCACACCTGG 2252
Qy 1223 CAGCGCAATTTCTCTCTGTCAACAATATCACATGGAATTAATCTTAAGGAAGTCTCTCTG 1282
Db 2253 TATCAGCTTCATGCTGGCGGAGAAATCTCCACAGGGGTGCTCTCTGGGAGAGGCCCG 2312
Qy 1283 TGCTTCAGTACTTGTACTATCTTGGTCAGATTGGTCTGGGAGATGTTGCCCATTTGAGCAACA 1342

```

Db      2313  TCTGCGAGTACCTGTATTACCTGGCTCAGATCGGCATGCCATGTCCCGCTCAGCAACA 2372
Qy      1343  ACTCCTATTCTTGTAGTATCATCGCAACCCCTTTTCCAAAGGTTCTTCCAAAGAGGTCTGA 1402
Db      2373  ACAGCCTGTTCCTTCAGTACACCGAACCCTCTCCCTGAGTACTTGTCCCGTGGCCTCA 2432
Qy      1403  ATGTCTCATTTATCTACGGATGACCCCTTGCAATTCACCTTGACAAAGAACCAATTCGTCG 1462
Db      2433  TGTGCTCGCTGTCCAGATGATCCCTTGAGTTCACATTCACCAAGGAGCCCTGATGG 2492
Qy      1463  AAGAATACAGCAATTCGTCTGCTGGAAGCTCAGTCTTCTGTGATTTATGCGAAATG 1522
Db      2493  AGGAGTACAGCATCGCCACCCAGGTGGAAGCTCAGCTCTGCGATATGTGCGAGCTGG 2552
Qy      1523  CGAGGAACCTGTTTACCAATCTGGGTTTTCATGCTCTCAAGGGCGCATGATGGTA 1582
Db      2553  CCGGTAAACGCGTCTCATGAGTGGCTTCTCTCAAGGTGAAAGCCATGCTGGGAC 2612
Qy      1583  AGAATCTACTTCAAGAGGACCTGCTGGAATGATATTCACAGAACCAATGTACCGACA 1642
Db      2613  CCAACTATACCAAGGAGGCCCTGAGGCGCATGATATCCGCGTACCAAGTCCAGACA 2672
Qy      1643  TCAGGGT 1649
Db      2673  TCCGAGT 2679

```

RESULT 15

```

SPADA1TCP
LOCUS      S. pombe adal and partial tcp1 genes.
DEFINITION X91498
ACCESSION  X91498
VERSION    1
KEYWORDS   adal gene; AMP deaminase; TCP1 gene.
SOURCE     Schizosaccharomyces pombe (fission yeast)
ORGANISM   Schizosaccharomycetes;
            Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
            Schizosaccharomycetales; Schizosaccharomycetaceae;
            Schizosaccharomycetes.
REFERENCE  1
AUTHORS    Rochet, M., Levesque, H. and Gaillardin, C.
TITLE      Putative AMP deaminase in S. pombe
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 4761)
AUTHORS    Nicaud, J.M.
TITLE      Direct Submission
JOURNAL    Submitted (14-SEP-1995) J.M. Nicaud, Institut National de la
            Recherche Agron., Lab. Genetique Moleculaire et Cellulaire,
            Thiverval-Grignon, 78850, FRANCE

```

FEATURES

```

source
1..4761
/organism="Schizosaccharomyces pombe"
/mol_type="genomic DNA"
/strain="972 h-"
/db_xref="taxon:4896"
/chromosome="II"
/map="left arm"
/clone="pMRSC1"
780..3083
/gene="adal"
780..3083
/gene="adal"
/codon_start=1
/product="AMP deaminase"
/protein_id="CAA62797.1"
/db_xref="GI:995562"
/db_xref="SWISS-PROT:P50998"
/translation="MNGTQNGRGLPSKSPRRSVSNASTVRSDDVNHKAGPGSKL
LNDLQSKIISIHMPHQGDNAVSVSGFETDPGNMETTFELFDEUAEIYLSIHK
CMDMRHKYISVQGLNDPDDSWIIYDPCKEGDDTGLFNFAKCKPGIENHEK
HMDHQIGFOYVENDSAYIAGTSPSHIPTIRDYIDLEFLLSASDGPFSFRILQY
LEGWNMYMLLNEYQELADTKVPHRDFYNVRKVDTHVHHSALANQKHLRLFIAKLR

```

gene

CDS

```

KCPNEKVIWRDGFELTQEVDFSLKLTYSYDLSIDTLDHHAHTDTFHRFDKFNLYKNI
GESSLRITFLATDNDINGRYLAELTKEVFTDLRTQKYQMAEYRISYIYGRNEEEDKLA
AWIDNELFSPFTVGLPKYLVMCMCIRSPVLRLLKRSEMSLNHSCFTKDPDRTHPK
LHVLQRIQVIGDSVDDSKSPERRTPRFPKPHNDINLNPYSYKLYMYANMISLNS
WRKIRGNFTFVLRPHCGEAGDTDLASAFSLSHGINHILLRKVPFLFYLYLWLOIPI
AMSPLSNNALFLAYDKNPFLLYFKRGLNLSLSTDDPLQFAFTREPLIEEYAAQAIYK
LSAVDMCELAENSVLQSGFERQLKRWLGVDFODIDRTNVP IIRLAYRALTLTQBIAL
NVKHVQSKHPSNHDLEELIHKYDAMTGTSDPLSASPRNTDATTSSRLSLHDGHDGA
RFPGLSVISERRRRKDSMASSQDLKD"
4456..4761
/gene="tcp1"
4456..4761
/gene="tcp1"
/codon_start=1
/protein_id="CAA62798.1"
/db_xref="GI:995563"
/db_xref="SWISS-PROT:P50999"
/translation="MSKAATVPVAFQDREKPOEVRLSNIMAARSVADAIKRTSLGPKGM
DKMIQTGKGEVILTNDGATILKHLVLPAAKMLVDLSAAQDVEAGDGTTSVVLGAS
"
```

BASE COUNT 1453 a 884 c 906 g 1518 t

ORIGIN

	Query Match	Best Local Similarity	Score	DB 8;	Length	4761;
Qy	205	GACTTGCAATATATTTCTCGAGTAACGGCTCGCAGGGAACACACAGAACATGTCGCCATAAT	25.0%;	DB 8;	Length	4761;
Db	1362	GATTTAGAAATTCCTTTCTCGGCTTCATCGGATGGTCTCTAGTAAATCATATTTCTTCGGT	60.1%;	Pred. No. 3.3e-89;		
Qy	265	CGGTTAAATCTCTTGGAGCATAAAGTTTAAATTTCCATCTGATGTTTAAATGCGGATAGGAA	0;	Mismatches	548;	Indels
Db	1422	AGATTGCGAGTATCTAGAAGGCGCTGGAAATATGATATGCTGCTAAACGAGTATCAAGAG	Conservative			
Qy	325	TTTCTTGCCGAGAGACTGCCCCACATCGTGAATTTTACAATGTGCGAAGGTTCACACT				
Db	1482	CTGGCAGATACCAAAAAAGTACCCCATCGTGATTTTACAATGTACGTAAAGGTAGATACT				
Qy	385	CATGTTTCATTCAGCATGCATGAATCAAAAAACATCTGTTGAGGTTTATAAAATCCAAA				
Db	1542	CACGTCCATCATTTCTGCTTTGCGAATCAAAAAACATCTCTTGCAGTTTATCAAGCTAAA				
Qy	445	CTAAGAAAAGAACCTGATGAGTGTCTATTTTCAGAGATGGTACTTATATGACTTTTAAAA				
Db	1602	CITTCGTAATGCCCTTAATGAAAAGTATCTGGCGTGATGGCAATTCCTGACTTTACNA				
Qy	505	GAGGTTTTTGAGAGCTTGGACTTAACTGGGTATGATCTGAAATGTTGATTTGCTAGATGTC				
Db	1662	GAAGTATTTGATTCGTTGAAGTTGACTTCGTATGACCTTAGTATAGACACGTTGGATATG				
Qy	565	CATGACAGAAAAGCACATTTTCATCGTTTGAACAATTCATCTTAAATACATCCATGTC				
Db	1722	CATGCTCATACTGACACCTTTTCCCGGATTTGACAAATTTAACTTGAATAATATCAAAAT				
Qy	625	GGCCAAAGTAGGCTCAGAGAAATTTTCTCAAAACAAGATAATCTTATTCAAGCGCGTTT				
Db	1782	GGTGAATCTCGTTTTCGGTACTATTTTCTGAAAACCGATATGATATTAACGGTCGATAT				
Qy	685	CTTGCTGAGTTGACAAAAGCAAGTTTCTCTGACCTTTCTGCTAGCAAAATATCAGATGGCA				
Db	1842	CITTGCTGAACTTACCAAGAGGTTTCCCGACCTTTAGAACTCAAAAGTACCAATGGCT				
Qy	745	GAATATAGGATTTCAATCTACGGAAGGAAACAGAGTGAATGGACCACTTCGAGTTGG				
Db	1902	GAATATCGTATTTCTATCTATGACGGAATCGTGAAGATGGGATAAGCTTCGACGCTGG				
Qy	805	ATAGTGAACAATGAATTGACACAGTGGAAATGTTGCTGGCTGGTTCAGATTCACACCTTA				
Db	1962	ATTATTGATACGAACCTTTTTTCT - CCAAGTTTCGTTGGCTTATTCAGTACTCTGTTG				
Qy	865	TATAATGTGTACAAGGAAATGGGTATCGTTACATCATTTCCAAAAATCTTCTTGACACAT				

Db 2021 TATGATGTATGAAGTCCGGTATTGTTGAGACTTTTGAAGAGGTGCTCAGAAATGTC 2080
Qy 925 TTGGTTCCCTTTTGAAGTT-ACTATTGATCCAGCTTCACACCCACAGCTCCATGTCTT 983
Db 2081 TTTGAACCACTGTTTGAAGTTTCAAAAGGATCTCGTACACATCTTAAATCCACGTGTT 2140
Qy 984 CTTGAAGCAGGTTGTAGGTTGGACCTGGTTGATGATGAAGTAAACAGAAAGGCTCC 1043
Db 2141 TTTGCAACGTGTTATCGGTTTGACTCTGTGACGACGAAAGTAAACCTGAAACGCCGTAC 2200
Qy 1044 AACAAAGCACATGCCACACACCTGAACAGTGGACCAATGTTCAACCCCTGCATTTTCATA 1103
Db 2201 CTTTCGCAATCCCTTATCCGAAACATTTGGGATATTAATTTAAATCCTCCATATAGTTA 2260
Qy 1104 TTATGCGTACTACTGTCTACTTAACCTATTACCCCTAAACAAAGCTCGTGAGTCAAGGG 1163
Db 2261 TTGGTTATATTACATGTATGCCAACATGACTAGCTTGAATAGCTGGAGAAAAATAGTGG 2320
Qy 1164 AATGACCACTATCAAAATTCGGTCCACATGCTGGAGAGGCTGGAGATGTTGATCAGTTGGC 1223
Db 2321 TTTTAACACCTTTGTCTTCTGCTCTCATTTGGTGAAGCTGGAGATACGGATCATTTGGC 2380
Qy 1224 AGCGACATTTCTCTCTGCACAAACATATCACATGGAATTAATCTAAGGAAGTCTCCTGT 1283
Db 2381 CAGTGTCTTTCTATTAAAGTACGGTATTAACCATGGTATTTTATTACGAAAGTTCCTTT 2440
Qy 1284 GCTTCAGTACTTGTACTACTTTGGTCAGATTTGGTCTGGCGATGTCCCCATTGAGCAACAA 1343
Db 2441 TCTTCAGTATCTTTGGTACCTTGATCAGATCCCTATCGCAATGTCTCCATTATCAACAA 2500
Qy 1344 CTCCTTATTTCTTGACTATCATCGCAACCCCTTTTCCAAAGTTCCTCCAAAGGCTGAA 1403
Db 2501 TGCACCTGTTTCTTTGCATATGATAAGAAATCCATTTTGAAGTACTTCAAACGAGGCTTAA 2560
Qy 1404 TGTCTCATTTATCTACGGATGACCCCTTTGCAATTCACCTGACAAAGAACCATTTGGTGA 1463
Db 2561 TGTAGTCTTTCCACTGATGATCCCTTTACAAATTTGCAATTTACCGAGAGCCTCTGATTGA 2620
Qy 1464 AGAATACAGCATTTGCTGCTGCTGTGGAAGCTCAGTTCTTTGTGATTTATGCGAAATTGC 1523
Db 2621 GGAATATGCTGAGCGCACAAATTTATAAGCTCTCTGCTGTCGATATGTGCGAATTGCG 2680
Qy 1524 GAGGAACCTGTTTACCAATCTGGGTTTTCATGCTCTCAAGGGCGCACTGGATTGGT 1581
Db 2681 TAGAAAACAGTGTTTTACAATCCGGTTTCGAACGTCAGCTTAAAGAACGTTGGTTGGGT 2738

Search completed: December 15, 2003, 20:04:24
Job time : 6853 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2003, 14:34:43 ; Search time 518 Seconds
(without alignments)
9953.526 Million cell updates/sec

Title: US-10-019-633-1

Perfect score: 1910

Sequence: 1 ccacggtccgggtgtgtcc.....aaaaaaaaaaaaaaaaaaaaa 1910

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1910	100.0	1910	22	AAF25488 Nucleotide sequenc
2	1836	96.1	2573	22	AAF25494 Nucleotide sequenc
3	1238	64.8	1988	22	AAF25497 Nucleotide sequenc
4	1101.4	57.7	2782	22	AAF25495 Nucleotide sequenc
5	1043.4	54.6	2520	24	AAF25496 A thaliana AMP dea
6	1041.8	54.5	2880	20	AAZ23395 A. thaliana AMP-de
7	1028	53.8	1803	24	ABZ19310 Arabidopsis thalia
8	962.4	50.4	2482	22	AAF25496 Nucleotide sequenc

9	959	50.2	1816	22	AAF25489	Nucleotide sequenc
10	531.6	27.8	914	24	AAF25493	A thaliana AMP dea
11	471.6	24.7	3386	24	ABK84112	Human cDNA differe
12	471.6	24.7	3386	24	ABK84112	Human benign prost
13	459.4	24.1	2250	23	ABK64551	Drosophila melanog
14	458.8	24.0	5442	22	ABK17219	cDNA encoding nove
15	458.8	24.0	5442	22	ABK64551	DNA encoding novel
16	447.8	23.4	2241	24	ABK64551	cDNA encoding rabb
17	341	17.9	1018	25	ABK54378	Aspergillus oryzae
18	282	14.8	282	24	ABK72727	Corn tassell-deriva
19	234	12.3	6405	23	ABK17218	Drosophila melanog
20	224.4	11.7	662	22	AAF25491	Nucleotide sequenc
21	211.6	11.1	551	22	AAF25490	Nucleotide sequenc
22	145.2	7.6	500	21	AAQ60866	Cat flea head and
23	111.8	5.9	322	14	AAQ60866	Human brain Expres
24	87.2	4.6	438	25	ABK38342	Bovine EST associa
25	87.2	4.6	438	25	ABK38342	Aspergillus oryzae
26	79.6	4.2	302	24	ABK25344	Human OREF polynuc
27	79.6	4.2	1248	21	AAZ6436	Human secreted pro
28	78.2	4.1	4531	23	ABV25119	Human prostate exp
29	77.4	4.1	388	22	AAH82121	Human polynucleoti
30	77.2	4.0	255	22	AAH82121	Rat differential t
31	77.2	4.0	1672	22	AAF92066	Human PRO1063 cDNA
32	77.2	4.0	1672	24	ABK74386	Human cDNA encodin
33	77.2	4.0	1672	25	ACA58818	cDNA encoding huma
34	77.2	4.0	1672	25	ACA60371	Novel human secret
35	77.2	4.0	1672	25	ACA63381	cDNA encoding huma
36	77.2	4.0	1701	21	AAZ64982	Membrane-bound pro
37	77.2	4.0	1701	22	AAZ64128	Human PRO4063 (UNQ
38	77.2	4.0	1701	25	ABK80196	Novel human secret
39	77.2	4.0	1701	25	ABK80700	Human secreted/tra
40	77.2	4.0	1701	25	ABK81083	Novel human secret
41	77.2	4.0	1701	25	ABK90173	Human secreted/tra
42	77.2	4.0	1701	25	ABK77784	Human PRO polynuc
43	77.2	4.0	1701	25	ABK79380	Human secreted/tra
44	77.2	4.0	1701	25	ABK54039	cDNA encoding huma
45	77.2	4.0	1701	25	ABK16983	Human PRO polynuc

ALIGNMENTS

RESULT 1
AAF25488
ID AAF25488 standard; DNA; 1910 BP.
XX
AC AAF25488;
XX
DT 15-MAY-2001 (first entry)
XX
DE Nucleotide sequence of a corn AMP deaminase enzyme.
XX
KW AMP deaminase; adenosine deaminase; adenosine; transgenic plant;
KW inosine; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 1..1737
FT FT /*tag= a
FT FT /product= "AMP deaminase"
XX
PN WO200109305-A2.
XX
PD 08-FEB-2001.
XX
PP 28-JUL-2000; 2000WO-US21009.
XX
PR 30-JUL-1999; 99US-0146473.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX (PION-) PIONEER HI-BRED INT INC.

Caspar T, Falco SC, Sakai H, Weng Z, Hu X;

WPI; 2001-159866/16.

P-PSDB; AAB31948.

New polynucleotides encoding AMP deaminase or adenosine deaminase used e.g. for designing or identifying herbicides that inhibit the enzyme activities, and as probes for genetic or physical mapping -

Claim 2; Page 42; 72pp; English.

The present sequence encodes an AMP deaminase. The specification also describes adenosine deaminase. These enzymes convert adenosine to inosine. Mutations in these genes cause disruptions in then salvage and catabolism of adenosine and AMP. In humans, this may lead to death of white blood cells, which causes severe immunodeficiencies. The AMP deaminase and adenosine deaminase may be used to prepare antibodies to these proteins, and to design or identify herbicides that inhibit their enzyme activities. The polynucleotides are used as probes for genetically and physically mapping genes that they compose, and as markers for traits linked to those genes, where such information may be used in plant breeding to develop lines with desired phenotypes. The nucleic acid fragments may be used to isolate cDNAs and genes encoding homologous proteins from the same or other plant species, and in polymerase chain reaction (PCR) protocols to amplify longer nucleic acid fragments encoding homologous genes from DNA or RNA. These may also be used to create transgenic plants in which the polypeptides are overexpressed or suppressed, and as probes in direct fluorescent in situ hybridisation (FISH).

Sequence 1910 BP; 596 A; 391 C; 383 G; 540 T; 0 other;

Query Match	Score 1910;	DB 22;	Length 1910;
100.0%	Score 1910;	DB 22;	Length 1910;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1910; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CCACGCGTCCGGGTGCTCCATGGGAGAGGGTCATAAATGACCCCTGTACTCCAAA	60
DB	1	CCACGCGTCCGGGTGCTCCATGGGAGAGGGTCATAAATGACCCCTGTACTCCAAA	60
QY	61	CTTAACCCCAACCGGTTCCACTTATGTGCTGAACCAAAAGTCAGAGCATGTTTTCAAAC	120
DB	61	CTTAACCCCAACCGGTTCCACTTATGTGCTGAACCAAAAGTCAGAGCATGTTTTCAAAC	120
QY	121	GTTGATGGCGTTATCCATGTTTATGCGGATAAAGATTGTAACGAGAGCAATTTACTCTGTG	180
DB	121	GTTGATGGCGTTATCCATGTTTATGCGGATAAAGATTGTAACGAGAGCAATTTACTCTGTG	180
QY	181	GCTGATGCTACAACTTCTTCACTGACTTGCAATTAATCTCCGAGTAACGGTGCAGGG	240
DB	181	GCTGATGCTACAACTTCTTCACTGACTTGCAATTAATCTCCGAGTAACGGTGCAGGG	240
QY	241	AACACAGAACCTGCTGCGTAATACGGTTAAATCTCTTCGAGCATTAAGTTTAAATTC	300
DB	241	AACACAGAACCTGCTGCGTAATACGGTTAAATCTCTTCGAGCATTAAGTTTAAATTC	300
QY	301	CTGATGTTTAAATGCGGATAGGAATTTCTTGCCACAGAAGACTGCCCCACATCGTGATTT	360
DB	301	CTGATGTTTAAATGCGGATAGGAATTTCTTGCCACAGAAGACTGCCCCACATCGTGATTT	360
QY	361	TACAATGTGCGAAGGTTTGACACTCATGTTTCATCATTCAGCATGCAATGATCAAAAACAT	420
DB	361	TACAATGTGCGAAGGTTTGACACTCATGTTTCATCATTCAGCATGCAATGATCAAAAACAT	420
QY	421	CTGTTGAGGTTTCATAAAATCCAAACTTAAGAAAAGAACCTGATGAGGTGGTCAATTCAGA	480
DB	421	CTGTTGAGGTTTCATAAAATCCAAACTTAAGAAAAGAACCTGATGAGGTGGTCAATTCAGA	480
QY	481	GATGGTACTTATATGACTTTTAAAGAGGTTTTTCAGAGCTTGACACTTAATCGGTATGAT	540
DB	481	GATGGTACTTATATGACTTTTAAAGAGGTTTTTCAGAGCTTGACACTTAATCGGTATGAT	540
QY	541	CTGAATGTGTTGATTTGCTAGATGTCCTATGCAGACAAAAAGCACATTTTCATCGTTTTGCAAA	600


```
Db 1578 GATTCACGCTTATATATGTTGTAAGAAATGGGTATCGTTTACATCATTCACAAATCT 1637
Qy 912 TCTTGACACATTTTCCTCTCTTTTTCAGGTTACTATTGATCCAGTTCACACCCACA 971
Db 1638 TCTTGACACATTTTCCTCTCTTTTTCAGGTTACTATTGATCCAGTTCACACCCACA 1697
Qy 972 GCTCCATGCTTCCTGTAAGCAGGTTGTAGGGTTGACCTGGTTGATGATGAAGTAAACC 1031
Db 1698 GCTCCATGCTTCCTGTAAGCAGGTTGTAGGGTTGACCTGGTTGATGATGAAGTAAACC 1757
Qy 1032 AGAAGGCTTCAACAAAGCAGATGCCACACACTGAACAGTGGACCAATGTGTTCAACCC 1091
Db 1758 AGAAGGCTTCAACAAAGCAGATGCCACACACTGAACAGTGGACCAATGTGTTCAACCC 1817
Qy 1092 TGCATTTTCATATTATGCGTACTACTGCTATGCTTAACCTTATTCACCCCTAAACAAGCTGG 1151
Db 1818 TGCATTTTCATATTATGCGTACTACTGCTATGCTTAACCTTATTCACCCCTAAACAAGCTGG 1877
Qy 1152 TGAGTCAAAAGGAATGACCACTATCAATTCGCTTCCACATGCTGGAGAGCTGGAGATGT 1211
Db 1878 TGAGTCAAAAGGAATGACCACTATCAATTCGCTTCCACATGCTGGAGAGCTGGAGATGT 1937
Qy 1212 TGATCATTTGGCAGCAGATTTCTTCTCTGTCAACATATCACATGGAATTAATCTAAG 1271
Db 1938 TGATCATTTGGCAGCAGATTTCTTCTCTGTCAACATATCACATGGAATTAATCTAAG 1997
Qy 1272 GAAGTCTCTGTGCTTCAGTACTTGTACTATCTTGTGTCAGATTGCTGCGCATGTCGCC 1331
Db 1998 GAAGTCTCTGTGCTTCAGTACTTGTACTATCTTGTGTCAGATTGCTGCGCATGTCGCC 2057
Qy 1332 ATTGAGCAACAACCTCTTATTTCTGACTATATCGCAACCCCTTTTCCAAAGTCTTCCA 1391
Db 2058 ATTGAGCAACAACCTCTTATTTCTGACTATATCGCAACCCCTTTTCCAAAGTCTTCCA 2117
Qy 1392 ACGAGTCTGAATGTCTCAATTAFTACGGATGACCCCTTTTGCAAATTCACCTGACAAAAGA 1451
Db 2118 ACGAGTCTGAATGTCTCAATTAFTACGGATGACCCCTTTTGCAAATTCACCTGACAAAAGA 2177
Qy 1452 ACCATTTGGTGAAGATACAGATTCGCTTCTGCTTGGAGCTCAGTCTCTTGATTT 1511
Db 2178 ACCATTTGGTGAAGATACAGATTCGCTTCTGCTTGGAGCTCAGTCTCTTGATTT 2237
Qy 1512 ATCGGAAATTCGAGGAACCTCTGTTTACCAATCTGGGTTTTCACATGCTCTCAAGGCGCA 1571
Db 2238 ATCGGAAATTCGAGGAACCTCTGTTTACCAATCTGGGTTTTCACATGCTCTCAAGGCGCA 2297
Qy 1572 CTGGATTGGTAAGAACTACTTCAAAAGAGACCTGCTGGAATGTATATTCACAGAACCAA 1631
Db 2298 CTGGATTGGTAAGAACTACTTCAAAAGAGACCTGCTGGAATGTATATTCACAGAACCAA 2357
Qy 1632 TGTACCGCATCAGGGTTCMAATTTAGAGAGATGATCTGGAGAAATGAATGAATGAACTAGT 1691
Db 2358 TGTACCGCATCAGGGTTCMAATTTAGAGAGATGATCTGGAGAAATGAATGAATGAACTAGT 2417
Qy 1692 GTACTCTGCAATGAGATCTTAATACACAGACGAGCTGGACCTGTAAGATGTCAGGCTCG 1751
Db 2418 GTACTCTGCAATGAGATCTTAATACACAGACGAGCTGGACCTGTAAGATGTCAGGCTCG 2477
Qy 1752 TGTATACAGACAGATTTGCTGTAGCTGTATGGAATPATCTTCACTTTTGTGTATG 1811
Db 2478 TGTATACAGACAGATTTGCTGTAGCTGTATGGAATPATCTTCACTTTTGTGTATG 2537
Qy 1812 CTTTCTCTATCTATGCAAAATTCACCTTGGAACTTC 1847
Db 2538 CTTTCTCTATCTATGCAAAATTCACCTTGGAACTTC 2573
```

RESULT 3

AAF25497

ID AAF25497 standard; DNA; 1988 BP.

XX

AAF25497;

XX

```
DT 15-MAY-2001 (first entry)
XX Nucleotide sequence of a soybean adenosine deaminase enzyme.
DE AMP deaminase; adenosine deaminase; adenosine; transgenic plant;
KW inosine; ss.
XX Glycine max.
XX Location/Qualifiers
FH Key 661..1697
FT CDS /tag= a
FT /product= "adenosine deaminase"
FT /transl_except= "(pos: 670..671, aa: Phe)"
XX WO200109305-A2.
XX 08-FEB-2001.
XX 28-JUL-2000; 2000WO-US21009.
XX 30-JUL-1999; 99US-0146473.
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX PA PIONEER HI-BRED INT INC.
XX Caspar T, Falco SC, Sakai H, Weng Z, Hu X;
XX WPI; 2001-159866/16.
XX P-PSDB; AAB31957.
XX New polynucleotides encoding AMP deaminase or adenosine deaminase used
XX e.g for designing or identifying herbicides that inhibit the enzyme
XX activities, and as probes for genetic or physical mapping
XX Claim 2; Page 65-66; 72pp; English.
XX The present sequence encodes an adenosine deaminase. The specification
XX also describes adenosine deaminase. These enzymes convert adenosine to
XX inosine. Mutations in these genes cause disruptions in then salvage and
XX catabolism of adenosine and AMP. In humans, this may lead to death of
XX white blood cells, which causes severe immunodeficiencies. The AMP
XX deaminase and adenosine deaminase may be used to prepare antibodies
XX to these proteins, and to design or identify herbicides that inhibit
XX their enzyme activities. The polynucleotides are used as probes for
XX genetically and physically mapping genes that they compose, and as
XX markers for traits linked to those genes, where such information may be
XX used in plant breeding to develop lines with desired phenotypes. The
XX nucleic acid fragments may be used to isolate cDNAs and genes encoding
XX homologous proteins from the same or other plant species, and in
XX polymerase chain reaction (PCR) protocols to amplify longer nucleic
XX acid fragments encoding homologous genes from DNA or RNA. These may
XX also be used to create transgenic plants in which the polypeptides are
XX overexpressed or suppressed, and as probes in direct fluorescent in
XX situ hybridisation (FISH).
XX Sequence 1988 BP; 574 A; 420 C; 429 G; 565 T; 0 other;
SQ
Query Match 64.8%; Score 1238; DB 22; Length 1988;
Best Local Similarity 86.4%; Pred. No. 2.3e-218;
Matches 1402; Conservative 0; Mismatches 215; Indels 5; Gaps 3;
Qy 122 TTGATGGCGTTATCCATGTTTATCGGATAAAGATTGTACGGAGACATTTATCTCTG-TG 180
Db 83 TGGCTGTTGTTCCCAAGTCTTTGTAGAATAAAGTCCGATCAAAAAGGATTATCTCTGTTT 142
Qy 181 GCTGATGCTACA-ACCTTCTTCACCTGACTTGATATATATTCCTCCGAGTAACGGCTGCAGG 239
Db 143 GCTGATGACGAGACCTTTTTCACCGACTTACATATGTTCTCCGGGTGACTGCCGCGG 202
Qy 240 GAACACAGAACTGTCTGCCATAATCGGGTTAAATCTTCTTGAGCATAAGTTTAAATCCA 299
Db 203 GAACACAGAACTGTCTGCCATAACCGATTGAATCTTCTTAGAACATCAAGTTCAAAATTTCA 262
```

QY 300 TCTGATGTTAAATGGGGATAGGGAATTTCTTCCGAGAGACTGCCCAACATCGTGATTT 359
DB 263 TCTGATGTTAAACGGGACAGGGAGTTCTTCCCAAAAACTGCACACATCGTGATTT 322
QY 360 TTACATGTCAGAAAGGTTGACACTGTTTATCATCTCAGCATGATGATCAAAACA 419
DB 323 TTACAAATGTTAGGAAGGTCGACACTCATGTTCCACACTCAGCATGATGATCAAAACA 382
QY 420 TCTGTTGAGGTTTCATAAAATCAAACTAAGAAAGAACTGATGAGGTGATTTTCAG 479
DB 383 TTTGCTGAGATTCATCAAGTCCAACTGAGAAAGAACTGATGAGGTGATTTTCAG 442
QY 480 AGATGGTACTTATATGACTTTAAAGAGGTTTTTGGAGCTTTGAGACTTTAACTGGGTATGA 539
DB 443 AGATGGTACTATATGACTTTTGAAGGAGGTTTTTGGAGCTTTGAGACTTTAACTGGGTATGA 502
QY 540 TCTGAATGTTGATTTGCTAGATGTCCTGTCAGTCCAGCAAAAGCAATTTTCATCGTTTGACAA 599
DB 503 CTGGAATGTTGATTTGCTAGATGTCCTGTCAGTCCAGCAAAAGCAATTTTCATCGTTTGACAA 562
QY 600 ATTCAATCTAAAATACAATCCCATGTCGCAAAAGTAGGCTCAGAGAAATTTTCTCATAACA 659
DB 563 ATTCACCTTAAATCAATCCATGTCGCAAAAGTAGGCTCAGAGAAATTTTCTCATAACA 622
QY 660 AGATAATCTTATTCAGGCGGTTTTTCTGCTGAGTTGACAAAGCAAGTTTTCTCTGACCT 719
DB 623 GGACAATCTTATTCAGGAGGATTTCTGCTGAGCTGACGATGCAAGTTTTTCTGACCT 682
QY 720 TTTCTGCTAGCAATATCAGATGGCAGATATAGGATTTCAATCTACGGAAGAAACAGAG 779
DB 683 TAATGCAAGCAAGTATCAGATGGCTGAATATAGGATTTTCAATCTATGGGAGGAACAAAG 742
QY 780 TGAATGGGACCAACTTGGCAAGTTGGATAGTGAACAAATGAATGCAAGTGGAAATGTTGT 839
DB 743 TGATGGGACCAACTTGGCAAGTTGGATAGTGAACAAATGAATGCAAGTGGAAATGTTGT 802
QY 840 CTGGCTGGTTCAGATTCACGCTTATATAATGTGTAACAGGAAATGGGTATCGTTACATC 899
DB 803 TTGGTTGATTCAGATTCACGCTTATATAACGTGTACAGCAAAATGGGCAATGTTACATC 862
QY 900 ATTCGAAATCTCTTGACACAAATTTTCTGCTGAGTTGAGTTACTTATGATCCAGC 959
DB 863 ATTTCAAAATCTCTTGACACAAATTTTCTGCTGAGTTGAGTTACTTATGATCCAGC 922
QY 960 TTCACACCCACAGCTCCATGTTCTTCTGAAGCAGGTTGTAGGTTGACCTGTTGATCA 1019
DB 923 TTGCGACCCGAGCTTCATGTTCTTCTAAAGCAGGTCGTAGGTTAGATTGTTGTTGATGA 982
QY 1020 TGAAGTAAACAGAAAGCGTCCAAACAAAGCACAATGCCCAACACCTGAA CAGTGGACCAA 1079
DB 983 TGAAGTAAACCTGAAAGCGTCCAACTAAGCACAATGCCCTACACCTGAAAGATGGAGAA 1042
QY 1080 TGTGTTCAACCTGCTGATTTTATATATATGCGTACTACTGCTATGCTAATTAATCACTCT 1139
DB 1043 TGTCTTCAACCCGGCAATTTTCAATATATGCTACTACTGCTATGCTAATTTGATACACAT 1102
QY 1140 AAACAGCTGCTGAGTCAAGGGAATGACCACTATCAAAATCCGTCACATGCTGGAGA 1199
DB 1103 GAACAAGCTGCTGAGTCAAGGGAATGATTAATCAAAATCCGTCACATGCTGGAGA 1162
QY 1200 GCGTGGAGATGTTGATCACTTGGCAGCAGCAATTTCTTCTGTCACAAATATCAATGATGG 1259
DB 1163 GCGTGGAGATGTTGATCACTTGGCAGCAGCAATTTCTTCTTGTGACAGTATATCAATGG 1222
QY 1260 AATTAATCAAGAAAGTCTCTGCTGCTTCAAGTACTTGTACTATCTTGTGTCAGATGGTCT 1319
DB 1223 AATCAATTTAAGAAAGTCTCTGCTGCTTCAATACCTGTACTACTCTTGTGTCAGATGGTCT 1282
QY 1320 GCGGATGTCCTCAGCAACAACTCTTATTTCTGACTATCATCGCAACCTTTCC 1379
DB 1283 GGAATGTCCTCAGCAACAACTCTTGTGTTGATTAACATCGGAACCTTTCC 1342

QY 1380 AACGTTCTTCAACGAGGTCTGAATGTCTCATTTATCTACGGATGACCCCTTTGCAAAATCA 1439
DB 1343 TATGTTTTTCAACGAGGACTGAATGTCTCGCTGTCCACGGATGATCCATTTGCAAAATCA 1402
QY 1440 CCTGACAAAACAACTTGGTGGAGAAATACAGCATTTGCTTCCGCTGTGGAAGCTCAG 1499
DB 1403 TCTGACAAAAGAGCAATTTGGTGGAGAAATACAGCATTTGCTTCCGCTGTGGAAGCTCAG 1462
QY 1500 TTTCTGTTGATTTATGCGAAATTTGCGAGGAATCTGTTTACCAATCTGGGTTTTTTCACATGC 1559
DB 1463 TTTCTGTTGATCTATGTGAATTTGCGAGAAATTTCTGTGTATCAATCAGGGTTTTTTCACATGC 1522
QY 1560 TCTCAGCGCACTGGATTGGTAAAGAACTACTTCAAAAGAGGACCTGCTGGAATGATAT 1619
DB 1523 TCTCAGGCACTTTGGATTGGCAAGAACTACTTCAAAAGAGGAGCCCTTTCAGGGAATGATAT 1582
QY 1620 TCACAGAACCAATGTACCGCACATCAGGGTTCAATTTTAGAGAGATGATCTCGAGAAATGA 1679
DB 1583 CCACAGAACGAATGTGCCACCCTCAGGATTTGAATTTAGGGACCTGATCTCGAGAGACGA 1642
QY 1680 AATGAACTAGTACTCTGCAATGAGATCTTAATACACGAGCTGGAACCTGTGTAAGA 1739
DB 1643 AATGAGCTCTCTTACCTCAACACGTCATCT---TGCTTACGAGGTGGACCAAGTAAGA 1699
QY 1740 TG 1741
DB 1700 GG 1701

RESULT 4
AAF25495
ID AAF25495 standard; DNA; 2782 BP.
AC AAF25495;
XX
XX
DT 15-MAY-2001 (first entry)
XX
DE Nucleotide sequence of a wheat AMP deaminase enzyme.
XX
KW AMP deaminase; adenosine deaminase; adenosine; transgenic plant;
XX inosine; ss.
XX
OS Triticum aestivum.
FH Key Location/Qualifiers
FT CDS 444..2489
FT /*tag= a
FT /product= "AMP deaminase"
XX
PN WO200109305-A2.
XX
XX
PD 08-FEB-2001.
XX
PF 28-JUL-2000; 2000WO-US21009.
XX
XX 30-JUL-1999; 99US-0146473.
XX
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Caspar T, Falco SC, Sakai H, Weng Z, Hu X;
XX
XX WPI; 2001-159866/16.
XX P-PSDB; AAB31955.
XX
XX New polynucleotides encoding AMP deaminase or adenosine deaminase used
XX e.g. for designing or identifying herbicides that inhibit the enzyme
XX activities, and as probes for genetic or physical mapping -
XX
XX Claim 2; Page 59-60; 72pp; English.
XX
XX The present sequence encodes an AMP deaminase. The specification also
XX describes adenosine deaminase. These enzymes convert adenosine to

QY 1513 TCGGAATTCGCGAGGAACCTCTGTTTACCAATCTGGGTTTTCACATGCTCTCAAGGCGAC 1572
 DB 2299 TCGGATAGCTCGTAACCTCAGTGTAACAGTCTGAGTTTCTACACGCCCTGAGTGGCAC 2358
 QY 1573 TGGATTGGTAAGAACTACTTTCAAAAGAGGACCTGCTGGAAATCATATTCACAGAACCAAT 1632
 DB 2359 TGGATTGGAAGAGATTACTACAAAGAGGACCTGATGGAAACGACATTCACAAACAAAC 2418
 QY 1633 GTACCCACATCAGGGTTCAATTTAGAGAGATGATCTGAGAGAAATGAATGAACCTAGTG 1692
 DB 2419 GTCCACACATAAGGCTGGAGTCCGTGACACGATCTGGAAGAGGAGATGCAACAGGTT 2478
 QY 1693 TACTCTGACAAATG 1705
 DB 2479 TATCTGGGCAAGG 2491

RESULT 6

AAZ23395
 ID AAZ23395 standard; DNA; 2880 BP.

XX AAZ23395;
 AC

XX 13-DEC-1999 (first entry)

XX A. thaliana AMP-deaminase DNA.

XX AMP-deaminase; adenosine triphosphate aminohydrolase; plant; herbicide;
 KW inhibitor; resistance; ss.

XX Arabidopsis thaliana.

XX Key Location/Qualifiers
 FT 78..2597
 CDS /*tag= a
 FT /product= "AMP-deaminase"

XX WO950400-A1.
 PN
 XX
 XX 07-OCT-1999.

XX 25-MAR-1999; 99WO-EP02016.

XX 01-APR-1998; 98DE-1014512.

XX (BADI) BASF AG.

XX Lerchl J, Reindl A;

XX WPI; 1999-580759/49.

XX P-PSDB; AAY33456.

XX DNA encoding AMP deaminase, assay systems for identifying inhibitors
 PT and transgenic plants -
 PS Claim 1; Fig 2; 45pp; German.

XX This invention describes a novel Arabidopsis thaliana AMP-deaminase
 CC (adenosine triphosphate aminohydrolase, EC 3.5.4.6) which has herbicidal
 CC activity. The AMP deaminase DNA sequence, can be introduced into pro- or
 CC eukaryotic cells, with the relevant control elements to control
 CC transcription and translation in the cell. An expression cassette
 CC derived from the products of the invention is used to transform plants.
 CC It is useful for producing a test system to identify inhibitors of AMP
 CC deaminase. Plants containing the expression cassette are useful for the
 CC production of the AMP deaminase. The plants have an increased resistance
 CC against inhibitors of the AMP-deaminase through strengthened expression
 CC of the AMP deaminase DNA sequence. The expression cassette is useful for
 CC producing plants with an increased content of AMP. This sequence encodes
 CC the A. thaliana AMP-deaminase described in the invention.

XX Sequence 2880 BP; 829 A; 600 C; 647 G; 804 T; 0 other;

Query Match 54.5%; Score 1041.8; DB 20; Length 2880;
 Best Local Similarity 76.0%; Pred No. 2.5e-182;
 Matches 1286; Conservative 0; Mismatches 407; Indels 0; Gaps 0;
 QY 13 GTTGCTCCATGGGAGAGAGGTCTATAAATGACCCCTGTCTCTCCAAAACCTAAACCCCAAC 72
 DB 876 GTTGACCATGGGAAAAGAGTCTATCTGATCTCTAGTCTCTCCAAAGCCTAAATACAGAG 935
 QY 73 CCGTTCACTTATGTCCTGAACCAAGTCAGAGCATGTTTCCAAACTGTTGATGCGGCTT 132
 DB 936 CCATTGACACTATCTCTCAGGAAAATCTGATCAATGTTTGTGAGATGCAAGATGGGTT 995
 QY 133 ATCCATGTTTATGCGGATAAAGATTGTACGGAGAGCAATTTATCTCTGCTGCTGATCTACA 192
 DB 996 GTCCACGTTGTTGCAATTAAGATGCAAAAGAGATCTCTTCCCGGTAGCTGATGCCACA 1055
 QY 193 ACCTTTCTTCACTGCTTGCATTATATCTCCGAGTAACGGCTGCAGGAAACAAAGAACT 252
 DB 1056 CGGTTTTTCACTGCTTGCATCAGTACTCAAAGTTCATAGCTGCAGGAAAACATCCGGACT 1115
 QY 253 GTCTCCATTAATCGGTTAAATCTTCTGAGCATAAGTTTAAATTCATCTGATGTTAAAT 312
 DB 1116 TTGTGCCACCGTCGACTAGTCTCTTAGAACAGAAAATTTAATCTCCATTTGATGCTTAAT 1175
 QY 313 GCGGATAGGGAATTTCTTCCCGAGAGACTGCCCCACATCGTGATTTTACAAATGTCAGG 372
 DB 1176 GCGGATAAAGAAATTTCTGCTCAAAAAGTGCACACATCGTGATTTTATTAACGTTAGG 1235
 QY 373 AAGTTGACACTCATGTTTCATCATTCAGCATGCGATGAATCAAAAACATCTGTTGAGGTT 432
 DB 1236 AAAGTCGACACTCATGTCATCAATTCAGCTTGCATGAACACAGAAACACCTTTTAAGTTT 1295
 QY 433 ATAAATCCAAACTAAGAAAGAACCTGATGAGGTGGTCATTTTCAGAGATGGTACTTAT 492
 DB 1296 ATTAAGTCAAAGCTCCGGAAGAACCCGATGAGGTTGTAATATATCCGAGATGGAACATAT 1355
 QY 493 ATGACTTTAAAGAGGTTTTCAGAGCTTGAGCTTAACTGGGTATGATCTGATGTTGAT 552
 DB 1356 TTGACCTTGAGAGAAGTTTTTGGAGGCTTGATCTGATGATATGACCTGAACTGCAC 1415
 QY 553 TTGCTAGATGTCATGACAGAAAAGCATTTCATCGTTTTTGAACAAATTCATCTAA 612
 DB 1416 CTTTGGGATGTTTCATGACAGAAAAGTACCTTTTCATCGTTTTTGAAGTTCAACCTAAAG 1475
 QY 613 TACAATCCATGTGGCCAAAGTAGGCTCAGAGAAATTTTCTCAACACAGATAATCTTAT 672
 DB 1476 TATAACCTTTGTGGTCAAAGTAGGCTTAGGGAGATTTTCTTAAACAGGATAATCTCATC 1535
 QY 673 CAAGGCCGTTTTTCTGCTGAGTTGACAAAGCAAGTTTTTCTCTGACCTTCTCTAGCAAA 732
 DB 1536 CAAGTCCGATTTCTTGGTGAGATAACAAGCAAGTCTTCTCTGACCTTGAAGCTAGTAA 1595
 QY 733 TATCAGATGGCAGAAATATAGGATTTCAATCTACGGAAGGAAACAGAGTGAATGGGACCAA 792
 DB 1596 TATCAGATGGCTGAATACAGATAATCTATATGTCAGAAAAAATGAGCGAGTGGACCAA 1655
 QY 793 CTTGCAAGTTGATAGTGAAACAAATGCAACAGTGAATGTTGCTGGCTGGTTTCAG 852
 DB 1656 CTCGCTAGTTGGATTTGTGAACAAATGATCTATACAGTGAAGATGTTGCTGGTTAAATTCAG 1715
 QY 853 ATTCCACGCTTATATAATGTGTACAGGAATGGGTATCGTTTACATCATTTCCAAAATCTT 912
 DB 1716 CTCACCGCTTGTACAAATTTACAGGACATGGGTATTTGACATCTGTTCCAGATATC 1775
 QY 913 CTTGCAACATTTTCGTTCTCTCTTTTGGAGTTTACTATTGATCCAGCTTTCACACCCACAG 972
 DB 1776 CTTGCAATATATTTCTTCTCTGTTTGAAGCCACGAGTAGATCTCTGATTTCCCTCTCAG 1835
 QY 973 CTTCAATGCTTCTCCTCAAGCAGGTTGTAGGTTGGACCTGGTTGATGATGAAGTAACCA 1032
 DB 1836 CTCCATGTTTTTTTGAAGCAGGTTGTTGATTTGATTTGATGATGATGATGATGATGATGAT 1895

QY 613 TACAATCCATGTGGCCAAAGTAGGCTCAGAGAAATTTCTCTCAAAACAGATATCTTATT 672
DB 700 TATAACCCCTGTGGTCAAGTAGGCTTAGGAGATTTCTCTTAAACAGGATATCTCATC 759
QY 673 CAAGGCGGTTTCTTCTGTGAGTTGACAAAGCAAGTTTCTCTGACCTTTCTGTGACAAA 732
DB 760 CAAGGTGCAATTTCTGTGAGATAACAAAGCAAGTCTCTCTGACCTTTGAAGCTAGTAAA 819
QY 733 TATCAGATGCGCAAAATATAGGATTTCAATCTCGGAGGAAACAGAGTGAATGGGACCAA 792
DB 820 TATCAGATGCGTGAATACAGAAATCTATATATGGCAGAAATATGACGAGTGGGACCAA 879
QY 793 CTTCGAAGTTGGATAGTGAACAAATGAATTCACAGTGGAAATGTTCTGTGGCTGGTTGAG 852
DB 880 CTGCGTAGTTGGATTTGGAACATGATCTATACAGTGAGATGTTCTGTGGTTAAATTCAG 939
QY 853 ATTCCAGCTTATATATATGTTGTAAGGAATGGGTATCGTTACATCATTTCCAAAATCTT 912
DB 940 CTCCACGCTTGTACAACTTTACAAAGGACATGGGTATTTGTGACATCGTTCCAGAAATATC 999
QY 913 CTTCGAACAAATTTCTGCTCTCTTTTGTAGGTTACTATTGATCCAGCTTCACACCCACAG 972
DB 1000 CTGGACAATATATTCATCTCTCTGTTTGAAGCACCGGTAGATCTGATCCCATCTCTCAG 1059
QY 973 CTCCATGCTCTCTGAGCAGAGTTGTAGGCTTGGACCTGTTGATGATGAAAGTAAACCA 1032
DB 1060 CTCCATGTTTGTGAGCAGGTTGTGGATTTGATTTGGTTCATGATGAACCAACCT 1119
QY 1033 GAAAGGCTTCAACAAAGCAGATGCCACACCTTGAAACAGTGGACCAATGTTTCAACCT 1092
DB 1120 GAAAGAGCTGCCAAACACATGCCACCTCCAGCTCAATGGACTAAACGATTCATCT 1179
QY 1093 GCATTTTCATATATGCTGCTACTGCTATGCTAACTTATTCACCTTCAACAGCTGCGT 1152
DB 1180 GCATTTTCGATATGCTGCTACTATGTTATGCTAACTCTATGTTTAAATAAGCTCGA 1239
QY 1153 GAGTCAAGGGAATGACCATCTCAAAATTCGCTCCACATGCTGGAGAGGCTGGAGATGTT 1212
DB 1240 GAGTCAAGGGAATGACCATCTCAAAATTCGCTCCACATGCTGGAGAGGCTGGAGATGTT 1299
QY 1213 GATCATTGGCAGGCAATTTCTCTCTGTGCAACATATCAATGAATTAATCTAAG 1272
DB 1300 GAGCATTGGCTGTACGTTTCTTAAACATGCTATGATGATGATGATGATGATGATGATGAT 1359
QY 1273 AGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1332
DB 1360 AGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1419
QY 1333 TTGAGCAACAACTCTTATTTCTGCTATCTATCTATCTATCTATCTATCTATCTATCTATCT 1392
DB 1420 CTGAGCAACAACTCTTATTTCTGCTATCTATCTATCTATCTATCTATCTATCTATCTAT 1479
QY 1393 CGAGGCTGAATGCTTCAATATCTATGATGATGATGATGATGATGATGATGATGATGATGAT 1452
DB 1480 AGAGGCTCAATGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1539
QY 1453 CCATTGGTGAAGATACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1512
DB 1540 CCTCTGCTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1599
QY 1513 TGCGAAATTCGAGGAACTCTGTTTCACTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1572
DB 1600 TGCGAGATAGCTGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1659
QY 1573 TGATTTGGTGAAGATCTCTTCAAAAGAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1632
DB 1660 TGGATTTGGAAGATTTACTTCAAAAGAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1719
QY 1633 GTACCGCACATCAGGTTTCAATTTAGAGATGATGATCTGGA 1672
DB 1720 GTGCCACATAGGGTGGAGTTCCGTTGACACCGGTATGGA 1759

RESULT 8
AAF25496
ID AAF25496 standard; DNA; 2482 BP.
XX AC AAF25496;
XX DT 15-MAY-2001 (first entry)
XX DE Nucleotide sequence of a corn adenosine deaminase enzyme.
XX AMP deaminase; adenosine deaminase; adenosine; transgenic plant;
KW inosine; ss.
XX OS Zea mays.
XX FH Key Location/Qualifiers
FT CDS 596..2482
FT /*tag= a
FT /product= "adenosine deaminase"
XX WO200109305-A2.
XX PN 08-FEB-2001.
XX PF 28-JUL-2000; 2000WO-US21009.
XX PR 30-JUL-1999; 99US-0146473.
XX PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX PI (PION-) PIONEER HI-BRED INT INC.
XX PI Caspar T, Falco SC, Sakai H, Weng Z, Hu X;
XX WPI; 2001-159866/16.
XX DR P-PSDB; AAB31956.
XX PT New polynucleotides encoding AMP deaminase or adenosine deaminase used
PT e.g. for designing or identifying herbicides that inhibit the enzyme
PT activities, and as probes for genetic or physical mapping
XX PS Claim 2; Page 63; 72pp; English.
XX CC The present sequence encodes an andenosine deaminase. The specification
CC also describes adenosine deaminase. These enzymes convert adenosine to
CC inosine. Mutations in these genes cause disruptions in then salvage and
CC catabolism of adenosine and AMP. In humans, this may lead to death of
CC white blood cells, which causes severe immunodeficiencies. The AMP
CC deaminase and adenosine deaminase may be used to prepare antibodies
CC to these proteins, and to design or identify herbicides that inhibit
CC their enzyme activities. The polynucleotides are used as probes for
CC genetically and physically mapping genes that they compose, and as
CC markers for traits linked to those genes, where such information may be
CC used in plant breeding to develop lines with desired phenotypes. The
CC nucleic acid fragments may be used to isolate cDNAs and genes encoding
CC homologous proteins from the same or other plant species, and in
CC polymerase chain reaction (PCR) protocols to amplify longer nucleic
CC acid fragments encoding homologous genes from DNA or RNA. These may
CC also be used to create transgenic plants in which the polypeptides are
CC overexpressed or suppressed, and as probes in direct fluorescent in
CC situ hybridisation (FISH).
XX SQ Sequence 2482 BP; 688 A; 539 C; 569 G; 686 T; 0 other;
Query Match 50.4%; Score 962.4; DB 22; Length 2482;
Best Local Similarity 78.7%; Pred. No. 9.5e-168;
Matches 1149; Conservative 0; Mismatches 311; Indels 0; Gaps 0;
QY 13 GTTGTCTCCATGGAGAGGAGGCTCAATAATGACCCCTGACTCCAAAACCTTAACCCCAAC 72
DB 1023 GTTGTCTCCGTTGGATTAAGAAGTTATATCCGACCCGACACACCCAGCTTAACCCAGAT 1082
QY 73 CCGTTCACTTATGTGCTGCTGAACCAAGTCAGAGCATGTTTTTCCAAACTGTTGATGCGGTT 132

1083 DB CAAATTTTATACATCTCTGAGGAAATTTCTGATCATATTTTGAATGCAAGATGGGTT 1142
1133 QY ATCCATGTTTATCGGATAAAGATTGTACGGAGAGCAATTTATCTCTGGCTGATGCTACA 192
1143 DB ATTGCTGATATCCAGATAGAGATGCAAAAGAGAGCTTTTCTCTAGCCGATGCAACT 1202
193 QY ACCTTCTTCACTGACTGATGATATATCTCGAGTAACGCTGACGGGAACAACAAGAACT 252
1203 DB ACATTTTTCACATGATCTTATCCTCACTTACTTCGAGTCATAGCAGCAGGGAATATAAGAACT 1262
253 QY GTCTGCATAATCGGTTAAATCTTCTTGAGCATAAAGTTTAAATTCATCTGATGTTAAAT 312
1263 DB TTATGCCATCATAGGCTCAATCTTCTAGAACAAAATTCATCTTCTATTTGATGCTAAAT 1322
313 QY GCGGATAGGGAATTTCTGCCAGAGACTGCCCAACATCGTGAATTTTCAATGTCAGG 372
1323 DB GCGGATAGAGAAATTTCTTCTCAGAGAGTGTCTCCACATCGAGACTTCTATAATGTTAGA 1382
373 QY AGGTTGACACTCATGTTCAATCTTCAATGATGAATCAAAAACATCTGTTGAGGTTTC 432
1383 DB AAGTTTGATGACTCATCTCCACCACTCAGCATGATGAATCAGAAACATCTTTTAAGGTTTC 1442
433 QY ATAAATCCAACTAAGAAAGAACTGATGAGTGGTCAATTTTCAGAGATGCTATTAT 492
1443 DB ATAAAGTCAAGCTGAGAAAGAGCCTGATGAGTTGTAATATTTTCGAGATGGGACATAT 1502
493 QY ATGACTTTTAAAGAGGTTTTTTCAGAGCTTTGGACTTAACCTGGGTGATGATCTGAATGTTGAT 552
1503 DB CTAAGTTTGAAGAGGTTTTTCAAGAGTTTAGATTTTGTCTGGATATGACCTCAATGTTGAC 1562
553 QY TTGCTAGATGCTCATGTCAGACAAAAGCAATTTTCATGCTGTTTTTGAACAAATCAATCTAAA 612
1563 DB CTTTGGAGCTTCACGCACAGAGTACTTTTTCATCGCTTTGATAAGTTTCAATCTTAA 1622
613 QY TACAATCCATGTCGCAAGTAGGCTCAGAGAAATTTTCTCAACAAAGATATCTTAT 672
1623 DB TACAATCCTTGGCTCAAGTAGGCTCAGGAGATATTTTCTTAAGCAGGATATCTCAT 1682
673 QY CAAGGCGTTTTCTGCTGAGTTTGAACAAAGCAAGTTTTTCTGACCTTTCTGCTAGCAAA 732
1683 DB CAAGGTCGTTTTCTTGGTGAGTTAACTAAGCAAGTTTTCAGATCTTCTGCTGCCAGTAA 1742
733 QY TATCAGATGGCAGATATAGGATTTCAATCTACGGAAGGAAACAGAGTGAATGGGACCAA 792
1743 DB TATCAGATGGCTGAATATAGAAATATCAATATATGTTAGGAAGCAAGTGAAGTGGACCAA 1802
793 QY CTTGCAAGTTGGATGATGACAAATGATTCACAGTGGAAATGTTGCTGGCTGGTTGAC 852
1803 DB CTAGCCAGTTGGATAGTGAATATGATTTGTACGCGGAATGTCGTTTGGTTGATTCAG 1862
853 QY ATTCACGCTTATATATGTTACAGGAATGGGTATCGTTACATCATTTCCAAATCTT 912
1863 DB CTTCCAGGTTGTACAAATGTTACAAAGAAATGGGAATTTGTACATCATTTCCAGAACATG 1922
913 QY CTTGCAACAACTTTCTGCTCTTTTGTAGGTTACTTATGATCCAGCTTTCACACCCACAG 972
1923 DB CTCGCAATATTTTCAATCCACTTTTGTAGGTCACGTCAACCCAGATTCACATCCTCAG 1982
973 QY CTCATGTTCTTCTGAGAGGTTGTAGGTTGGACCTGTTGATGATGAAGTAACCA 1032
1983 DB CTGCAATGTTTTCTGAAACAGGTTGTGTTGGTTGGATTTGTTGATGATGAAGCAACCT 2042
1033 QY GAAAGGCGTCCAAACAGACATGCCACACCTGAAACAGTGGACCAATGTTTCAACCT 1092
2043 DB GAAAGGCGCCAAACAAACACATGCTACATCTGACCAATGACATATGTTTCAATCG 2102
1093 QY GCATTTTCATATATGCGTACTGCTATGCTAACTTATTCACCTTAAACAGCTGCT 1152
2103 DB GCATTTTCATATGCTATGCTATGCTATGCTAACTTATTCACCTTAAACAGCTTGA 2162
1153 QY GAGTCNAAGGGAATGACCACTATCAATTTCCGTCACATGCTGGAGAGCTGGAATGTT 1212

2163 DB GAATCAAGGGAATGACAAACATCAAAATTCGTCACATTTCTGGAGAGCTGGTGATATT 2222
1213 QY GATCACTTTGGCAGGACATTTTCTCTCTGTCAACAATATCACATGGAATTAATCTAAGG 1272
2223 DB GACCACCTTGGCAGCAACCTTTCTCAGGCTCACAACATTTGCACATGGAATCAATTTGAAA 2282
1273 QY AAGTCTCCTGCTCAGTACTTGTACTATCTTGTGATGATGCTGCGGATGTCCTCA 1332
2283 DB AAATCTCCTGCTTCAATATATTTATTTAGCCAGATTTGGGCTGGCAATGTCTCT 2342
1333 QY TTGACCAACAACTCCTTATTTCTTGACTATCATCGAACCCCTTTTCCAACTGTTCTTCAA 1392
2343 DB TTGACATAACTCCCTTATTTCTAGACTACATCGGAATCTTTTCCAACTGTTCTCTTA 2402
1393 QY CGAGTCTGAATGTCTCATTTCTACGGATGACCCCTTTTGCAAATTCACCTGACAAAGAA 1452
2403 DB CGGGTCTGAATGTGTCACTTTCTACTGATGATCTCTCCAAATTCACCTTAAACAAAGAA 2462
1453 QY CCATTGGTGGAGATACAG 1472
2463 DB CCATTGGTGAAGATATAG 2482

RESULT 9
AAF25489
ID AAF25489 standard; DNA; 1816 BP.
XX AAF25489;
AC AAF25489;
DT 15-MAY-2001 (first entry)
XX Nucleotide sequence of a rice AMP deaminase enzyme.
DE AMP deaminase; adenosine deaminase; adenosine; transgenic plant;
KW inosine; ss.
XX Oryza sativa.
OS
XX
PH Key Location/Qualifiers
FT 3..1481
CDS /*tag= a
FT /product= "AMP deaminase"
XX
PN WO200109305-A2.
XX
PD 08-FEB-2001.
XX
PF 28-JUL-2000; 2000WO-US21009.
XX
PR 30-JUL-1999; 99US-0146473.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX (PION-) PIONEER HI-BRED INT INC.
XX
PI Caspar T, Falco SC, Sakai H, Weng Z, Hu X;
XX
XX WPI; 2001-159866/16.
DR P-PSDB; AAB31949.
XX
XX New polynucleotides encoding AMP deaminase or adenosine deaminase used
PT e.g. for designing or identifying herbicides that inhibit the enzyme
PT activities, and as probes for genetic or physical mapping -
XX
PS Claim 2; Page 44-45; 72pp; English.
XX
CC The present sequence encodes an AMP deaminase. The specification also
CC describes adenosine deaminase. These enzymes convert adenosine to
CC inosine. Mutations in these genes cause disruptions in then salvage and
CC catabolism of adenosine and AMP. In humans, this may lead to death of
CC white blood cells, which causes severe immunodeficiencies. The AMP
CC deaminase and adenosine deaminase may be used to prepare antibodies
CC to these proteins, and to design or identify herbicides that inhibit
CC their enzyme activities. The polynucleotides are used as probes for

CC Genetically and physically mapping genes that they compose, and as
 CC markers for traits linked to these genes, where such information may be
 CC used in plant breeding to develop lines with desired phenotypes. The
 CC nucleic acid fragments may be used to isolate cDNAs and genes encoding
 CC homologous proteins from the same or other plant species, and in
 CC polymerase chain reaction (PCR) protocols to amplify longer nucleic
 CC acid fragments encoding homologous genes from DNA or RNA. These may
 CC also be used to create transgenic plants in which the polypeptides are
 CC overexpressed or suppressed, and as probes in direct fluorescent in
 CC situ hybridisation (FISH).

XX Sequence 1816 BP; 566 A; 344 C; 381 G; 525 T; 0 other;

Query Match 50.2%; Score 959; DB 22; Length 1816;
 Best Local Similarity 79.4%; Pred. No. 3 9e-167;
 Matches 1136; Conservative 0; Mismatches 295; Indels 0; Gaps 0;

QY 261 TAATCGGTTAAATCTTCTTTGAGCATAAAGTTTAAATCCATCTGATGTTAAATGCGGATAG 320
 DB 8 TAAACGTTTAAATCTTCTAGACAGAAATTCATCTTCAATTTGATGGTCAATGCCGATAG 67
 QY 321 GGAATTTCTTCCAGAGAGCTGCCACATCTGATTTTACAAATGTCAGGAGTTGA 380
 DB 68 AGAACTACTTGTCTGAGAAAGCTGCACCCATCGGAGCTTCTACAATGTCAGGAGTTGA 127
 QY 381 CACTCATGTTTCATCATCTCAGCATGCAATGAATCAAAACATCTCTGAGGTTTCATAAAATC 440
 DB 128 TACTCATGTTTCATCTCAGCATGCAATGAATCAAAACATCTCTGAGGTTTCATAAAATC 487
 QY 441 CAAACTAAGAAAGAACCTGATGAGTGGTTCATTTTTCAGAGATGGTACTTATATGACTTT 500
 DB 188 CAAAGTTGAGAAAGAACCTGACGAGGTTGTGATTTTATAGAGATGGTACTTATGACTCT 247
 QY 501 AAGAGAGTTTTCAGAGCTTGAATTCATCGTTTTCAGAAATTCATCTGAGTTCATGTTGCTAGA 560
 DB 248 TAAGGAGGTTTTCAGAGCTTGAATTCATCGTTTTCAGAAATTCATCTGAGTTCATGTTGCTAGA 567
 QY 561 TGTCCATGACAGAAAGACATTTTCATCGTTTTCAGAAATTCATCTGAGTTCATGTTGCTAGA 620
 DB 308 TGTGATGCGATTAAGATACATTCATCGTTTTCAGAAATTCATCTGAGTTCATGTTGCTAGA 367
 QY 621 ATGTGCGCAAGTAGGCTCAGAGAAATTTTCTCAAAACAGATAATCTTATTCAGAGCGG 680
 DB 368 TTGTGCGCAATCCCGCTGAGGAGATCTTTCTTAAACAGGACAACTTATTCAGAGCGG 427
 QY 681 TTTTCTGCTGAGTTGACAAAGCAAGTTTCTCTGACCTTTCTGCTAGCAATATCAGAT 740
 DB 428 ATTCTTGTCTGAAATGACAAAGAAAGTATTTCTGATCTTGAAGCAAGTAAATATCAGAT 487
 QY 741 GGCAGAAATATAGGATTTCAATCTACGAGGAAACAGAGTGAATGGACCACTTCGAAG 800
 DB 488 GGTGATGATAGATATCTATCTATGGGAAAGAAAGATGAGTGGATCAGATGCGAAG 547
 QY 801 TTGGATAGTGAACAAATGAATTCACAGTGAATGTTGTCTGCTGCTTCAAGATTCACG 860
 DB 548 CTGGATAGTGAATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 607
 QY 861 CTTATATATGCTACAGAAATGGTATCGTTTACATCATCTTCAAAATCTTTCACAA 920
 DB 608 GATATACATGATATACAGGAGATGGAAACAAATCTTTCAGAACTCTTTCACAA 667
 QY 921 CATTTTCTGTTCTTCTTTCAGGTTTACTATTGATCCAGCTTTCACACCAAGCTCCATGT 980
 DB 668 TATTTTCTGCTTCTTTCAGGTTTACTATTGATCCAGCTTTCACACCAAGCTCCATGT 727
 QY 981 CTTCTGAGACGAGTTGTAGGTTTGGACCTGTTGATGATGATGATGATGATGATGATGATGATGAT 1040
 DB 728 TTTCTTTCGACAGGTCGTTGGGCTGGATTTAGTGGATGATGATGATGATGATGATGATGATGATGAT 787
 QY 1041 TCCAAACAGACATGCCCACTGACAGTGGACCAATGTTTCAACCTGCAATTTTC 1100
 DB 788 CCCAAACAAACACATGCCCTACCTGAGCAATGACTAATGTTTTCATTCAGCATATGC 847

QY 1101 ATATTATCGTACTACTCTGCTATGCTAACTATTTCACCCCTAAACAGCTGCGTGAAGTCAAA 1160
 DB 848 ATATTATGCTACTATTGTTATGCTAACTTTGACAGCTGAAACAGCTTTCGTTGAGTCCAA 907
 QY 1161 GGGATATGACACTATCAAAATTCGCTCCACATGCTGGAGAGCTGGAGATGTTGATCACTT 1220
 DB 908 GGGTATGACAAACAAATCAAACTTCGTTCCACACTGTTGGGAGGCTGGAGATATTTGATCATCT 967
 QY 1221 GGCAGCGACATTTCTTCTCTGTCACAAACATATACATGGAATTAATCTAAAGAACTCTCC 1280
 DB 968 TGCTGCAGCATTTCTTACTTCTCAATAATATGCTCAGGGGTTAATTTAAAGAACTCTCC 1027
 QY 1281 TGTGCTTCAGTACTTGTACTATCTTTGTCAGATGCTGCGAGTGTCCCACTTGAGCAA 1340
 DB 1028 TGTCTCTCAGTATCTGTTATACCTAGCTCAGATGCTTGTCTTGCCATGTTCTCTTTGAGCAA 1087
 QY 1341 CAACTCTCTTATTTCTTGACTATCATCGCAACCCCTTTTCCAAAGTTCTTCCAAAGAGGCT 1400
 DB 1088 CAACTCAATGTTTATTTGATTTATCACGAAACCCCTTTTCCCAACATTTTCTTAAGAGGCT 1147
 QY 1401 GAATGCTCATTTATCTACGATGACCCCTTTCGAAATTCACCTGACAAAGAACCACTTGT 1460
 DB 1148 TAACTTTCTCTATCAACGATGACCCCTTTCGAAATTCACCTGACAAAGAACCACTTGT 1207
 QY 1461 GGAAGATACAGCATTTGCTGCTTCTGCTGGAAGCTCAGTTCTTGTGATTTATGCGAAAT 1520
 DB 1208 TGAAGAAATATAGCATCGCTGCTGCTGTTGAAAGCTTAAGTTTCATCGACCTATGTAAT 1267
 QY 1521 TCGGAGGAATCTGTTTACCAATCTGGGTTTTCACATGCTCTCAAGGCGCACTGATGG 1580
 DB 1268 TGCTAGGAATTTCTGTTGACCACTGCTGTTTCTCTCATAGGCTCAAGTCACTGATGG 1327
 QY 1581 TAAGAACTACTTCAAAAGAGGACCTGCTGGAATTCATATTCACAAACCAATGTACCGCA 1640
 DB 1328 GAGAACTACTACAAAGAGGTCATGTCGCAATGACATTCACAGACAAATGTTCTCTCA 1387
 QY 1641 CATCAGGTTCAATTTAGAGAGATGATCTGGAGAAATGAAATGAAATAGT 1691
 DB 1388 CATCAGGATGAAATTCGACACACTATTTTGGAAAGAAATGGAGCTAAT 1438

RESULT 10

AAL45031

ID AAL45031 standard; cDNA; 914 BP.

XX AAL45031;

AC AAL45031;

XX 16-MAY-2002 (first entry)

DT 16-MAY-2002 (first entry)

XX A thaliana AMP deaminase coding sequence fragment EST.

DE AMP deaminase; adenosine monophosphate; transgenic plant; EST;

KW herbicide resistance; herbicide; inhibitor; expressed sequence tag; ss.

KW Arabidopsis thaliana.

OS Key Location/Qualifiers

XX CDS 2..880

FT /*tag= a

FT /product= "AMP deaminase fragment"

XX WO200206319-A2.

XX 24-JAN-2002.

XX 06-JUL-2001; 2001WO-EP07767.

XX 17-JUL-2000; 2000DE-1035084.

XX (AVET) AVENTIS CROPS SCIENCE GMBH.

XX Schulz A, Streiber W, Hanke C, Schmidt F, Schubel A;

XX

CC inflammation with an agent that modulates the expression of gene(s)
CC from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for
CC modulating GA; M3 is useful for screening an agent capable of modulating
CC GCA preferably in an inflammation in a tissue; M4 is useful for
CC detecting an inflammation (especially chronic) in a tissue, an allergic
CC response in a subject, exposure of a subject to a pathogen or sterile
CC inflammatory disease (e.g. psoriasis, rheumatoid arthritis,
CC glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal
CC reperfusion injury, ARDS, adult respiratory distress syndrome,
CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,
CC periodontal disease; also bacterial infection, viral infection,
CC parasitic infection, protozoal infection, fungal infection and M5 is
CC useful for treating one of the above conditions. The present
CC sequence represents a gene differentially expressed in granulocytes.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX

SQ Sequence 3386 BP; 658 A; 1046 G; 973 G; 709 T; 0 other;

Query Match 24.7%; Score 471.6; DB 24; Length 3386;

Best Local Similarity 58.3%; Pred. No. 1.3e-77;

Matches 847; Conservative 0; Mismatches 604; Indels 3; Gaps 1;

QY	238	GGGAACACAGAACTCTGCCATAATCGTTAAATCTTCTTGAGCATAAAGTTTAATTC	297
DB	956	GGCCCCATAAGTCATTCTGCTACCGCGCTGCAGTACTGAGCTTCAAGTTCCAGATG	1015
QY	298	CATCTGATGTTAATCGGATAGGAAATTTCTTGGCCAGAAGACTGCCCCACATCGTAT	357
DB	1016	CATGCTCTACTCAATGAGATGAAGAGCTGGCGCCAGAAGAGTGCACACCCAGAT	1075
QY	358	TTTTACAATGTCAGGAAGTTGACATCATGTTTCATCATTCACGATGATGATCAAAA	417
DB	1076	TTCTACAACATCCGCAAGGTGGACACCCACATCCATCCCTGCTGTCATGAACCAAG	1135
QY	418	CATCTGTGAGGTTCATAAAATCCAACTAAGAAAAGAACTCATGAGGTGGTCAATTTTC	477
DB	1136	CATCTGCTGGCTTCATCAAGCGGGCANTGAAGCGGCACCTGGAGAGATCTGTGACGTG	1195
QY	478	AGAGATGGTACTTATATGACTTTTAAAGAGGTTTTTGGAGCTTGGACTTAACTGGGTAT	537
DB	1196	GAGCAGGGCGTCAACAGACGCTGCGGAGGTCTTTTGAGAGCATGAATCTCACGGCTAC	1255
QY	538	GATCTGAATGTTGATTTGCTAGATGTCATGACAGACAAAGACATTTTCATGTTTGAC	597
DB	1256	GACCTGAGTGTGACACGCTGGATGTGCATGCGGACAGGAACACTTTTCCATGCTTTGAC	1315
QY	598	AAATTCATCTAAATACAAATCCATGTGGCCAAAGTAGGCTCAGAGAAATTTTCTCAAA	657
DB	1316	AAAGTTTAATGCCAATACAAACCTATTTGGGAGTCCGTCTCCAGAGATCTTTCATCAAG	1375
QY	658	CAAGATAATCTTATCAAGSCCGTTTTCTTGTGCTGAGTTGACAAAGCAAGTTTTCTCTGAC	717
DB	1376	ACGGACAACAGGTAATCTGGGAAGTACTTTTGTCTCACATCATCAAGGAGTGTGTGAC	1435
QY	718	CTTTCTGTAGCAAAATATCAGATGGCAGATATAGGATTTCAATCTACGGAAGAAACAG	777
DB	1436	CTGGAGGAGACCAATACCAAGATGCAGAGCTCGGCTCTCCATTTACGGGCGCTCGAGG	1495
QY	778	AGTGAATGGGACCAACTTGAAGTTGGATAGTGAACAATGAATTTGACAGTGAAGATGTT	837
DB	1496	GATGAGTGGGACAAAGCTGGCGCGCTGCGCGCTCATGACCGCGTGCATCCCCCAACGTG	1555
QY	838	GTCTGGCTGGTTCAGATTTCCAGCTTTATATAATGTGTACAGGAATGGGTATCGTTACA	897
DB	1556	CGCTGGCTGGTTCAGATGGCTGGCGCGCTGCGCGCTCATGACCGCGTGCATCCCCCAACGTG	1615
QY	898	TCATTCGAAATCTTCTTGACAAATTTTGTCTCTTTTGGAGTTACTATTGATCCA	957
DB	1616	AACTTCCAGGATGCTGGGAACATCTTCTCCCACTGTTCGAGGCCACTGTGCACCTT	1675

QY	958	GCTTCACACCCACAGCTCCATGTCTTCTGAAGCAGGTTGTAGGTTGGACCTGTTGAT	1017
DB	1676	GCCACCCACCGGAATCTCATCTCTTCTTAGCAGCAGTGGATGTTTTCAGACGCTGGAT	1735
QY	1018	GATGAAGTAAACGAAAGGGTCCAAACAAGC---ACATGCCACACCTGGAACAGTGG	1074
DB	1736	GATGAGTCCAAAGCTGAAACCATGTCTTCAACCTGGAGAGCCCTCCCTGAGCGGTGG	1795
QY	1075	ACCAATGTTTCAACCTGCAATTTTATATTATGCTACTACTGTCTGCTAACTTATTC	1134
DB	1796	GTGGAGGAGGACAAACCCATGCTTACCTGTACTACCTTGTCCACATGAGCC	1855
QY	1135	ACCCTAAACAAAGCTGCTGAGTCAAGGGAATGACCACTATCAAAATCCGTCCACATGCT	1194
DB	1856	ATGTTGAACCACTGCGCAGGAGGAGGCTTCCACAGCTTTGTCTGAGGCCACACTGT	1915
QY	1195	GGAGAGCTGGAGATGTTGATCATTGTCAGGAGGACATTTCTCTGTCTGACACATATCA	1254
DB	1916	GGGAGGCTGGGCCCATCCACCATGCTGTGTGAGCTTTCATGCTGGCTGAGAACATTTCC	1975
QY	1255	CATGGAATTAATCTAAGGAAGTCTCTGCTTTCAGTACTTGTACTATCTTGGTCCAGATT	1314
DB	1976	CACGGGCTCTTCTGCGAAGGCCCTCTGCTGAGTACTGCTACTTCTGAGCCAGATC	2035
QY	1315	GGTCTGGGATGTCCCAATTTGAGCAACAACTCTTATTTTGTGACTATCATCGCAACCT	1374
DB	2036	GGCATCGCATGTCTCCGCTCAGCAACAAAGCCCTTCTCCTCAGCTATCACCGGAATCG	2095
QY	1375	TTTCCAAAGTCTTCCAAAGGAGTCTGAATGTCTCATTTATCTACGGATGACCTTTGCA	1434
DB	2096	CTACCGGAGTACTCTGCGCGGCTCATGTGTCTCCCTGTCCACTGATGATCCCTTGCAG	2155
QY	1435	ATTCACTGCAAAAGAACCATTTGGTGAAGAATACAGCATTTGCTGCTTGGTGGGAAG	1494
DB	2156	TTCCACTTCAACAGAGCCGCTGATGAGGAGTACAGCATGCCACCCAGGTGTGGAG	2215
QY	1495	CTCAGTCTTGTGATTTATGCGAAATTCGAGGAACTCTGTTTACCAATCTGGGTTTTCA	1554
DB	2216	CTCAGCTCTGCGATATGTGTGAGCTGGCCGCAACAGCGTCTCATGAGCGGCTTCTCG	2275
QY	1555	CATGCTCTCAAGCGGCTGAGTGGTGAAGAATCTACTTCAAAAGAGGACCTGCTGGAAT	1614
DB	2276	CACAAGGTAAGAGGACCTGCTGGGAGCCCAACTATACCAAGGAAGGCCCTGAGGGAAAT	2335
QY	1615	GATATTCACAGAACCAATGTACCGCACATCAGGGTTCAATTTAGAGAGATGATCTGGAGA	1674
DB	2336	GATCCCGCGGACCAATGTGCCAGATCCCGTGGGTACCGCTACGAGACCTGTGTC	2395
QY	1675	AATGAATGAACCT	1688
DB	2396	CAGGAGCTGGCGCT	2409

RESULT 12

ABK64551

ID ABK64551 standard; DNA; 3386 BP.

XX

AC ABK64551;

XX

DT 18-JUN-2002 (first entry)

XX

XX Human benign prostatic hyperplasia gene #446.

XX Human; benign prostatic hyperplasia; BPH; prostate cancer; gene; ds.

XX Homo sapiens.

XX OS

XX WO200212440-A2.

XX PN

XX 14-FEB-2002.

XX PD

XX 07-AUG-2001; 2001WO-US24708.

XX PF

XX XX

PR 07-AUG-2000; 2000US-223323P.
 PR 05-JUN-2001; 2001US-0873319.
 XX (GENE-) GENE LOGIC INC.
 PA (NIBS) JAPAN TOBACCO INC.
 XX Munger WE, Kulkarni P, Getzenberg RH, Waga I, Yamamoto J;
 XX WPI; 2002-257476/30.
 XX
 XX Identifying drugs for and diagnosing benign prostatic hyperplasia, by
 PT detecting expression levels of one or more genes in prostate cells from
 PT patient that are differentially regulated compared to normal prostate
 PT cells -
 XX
 XX Disclosure; Page 262-263; 444pp; English.
 XX
 CC The invention relates to a method of diagnosing (I) the onset or
 CC progression of benign prostatic hyperplasia (BPH), or screening (II) for
 CC or identifying an agent that modulates the onset or progression of BPH.
 CC The method is based on changes in gene expression in BPH tissue isolated
 CC from patients exhibiting different clinical states of prostate
 CC hyperplasia as compared to normal prostate tissue. (I) comprises
 CC detecting the expression levels of one or more genes in prostate cells
 CC from the subject that are differentially regulated compared to normal
 CC prostate cells. (II) comprises preparing a first gene expression profile
 CC of BPH cells or BPH-like cell population, exposing the cells to the
 CC agent, preparing a second gene expression profile of the agent exposed
 CC cells, and comparing the first and second gene expression profiles.
 CC (I) is useful for diagnosing the onset or progression of BPH. (II) is
 CC useful for identifying an agent that modulates the onset or progression
 CC of BPH. The methods are useful to present information identifying
 CC the expression level in a tissue or cells, by comparing the expression
 CC level of genes given in the specification in the tissue or cells to the
 CC level of expression of gene in the database, and displaying the
 CC expression levels of at least one gene in the tissue or cell sample
 CC compared to the expression level in BPH. Agents using (II) are useful for
 CC treating BPH or prostate cancer. ABK64106-ABK64860 represent human
 CC benign prostatic hyperplasia gene sequences of the invention.
 XX
 SQ Sequence 3386 BP; 658 A; 1046 C; 973 G; 709 T; 0 other;
 Query Match 24.7%; Score 471.6; DB 24; Length 3386;
 Best Local Similarity 58.3%; Pred. No. 1.3e-77;
 Matches 847; Conservative 0; Mismatches 604; Indels 3; Gaps 1;
 QY 238 GGGACACAGAACTGCTGCCATAATCGTTAAATCTTTTGAGCATAAGTTTAAATTC 297
 DB 956 GGGCCCAATAAAGTCATCTGCTACCGCCGGCTGCAGTACCTGAGCTCCAAGTTCAGATG 1015
 QY 298 CATCTGATGTTAAATGCGGATAGGGAATTTCTTGCCAGAGAAGTGCCTCCACATGCTGAT 357
 DB 1016 CATGTCTACTCAATGAGATGAAGAGCTGGCGCCAGAGAAAGTGCACACCGAGAT 1075
 QY 358 TTTTACAATGTCAGGAGTTGACACTCATGTTTCATCTATTCATGACATGATGATCAAAA 417
 DB 1076 TTCTACAACATCGCAAGGTGACACCCACATCCATGCTCGCTCTGATGAACACAGAG 1135
 QY 418 CATCTGTTGAGTTTCAATAATCCAACTAAGAAAGAACCTGATGAGTGTCTATTTTC 477
 DB 1136 CATCTGCTCGGTTTCAATGAGGGGCAATGAAGCGGCACCTGGAGGATCTGTCAGCTG 1195
 QY 478 AGAGATGTTGACTATATGACTTTAAAGAGGTTTTTGGAGGCTTGGACTTAATCTGGGTAT 537
 DB 1196 GAGCAGGGCCGGAACAGACGCTGGGAGGTTCTTGAGAGCATGAATCTACGGCTTAC 1255
 QY 538 GATCTGATGTTGATTTGCTAGATGTCATGTCAGACAAAGACATTTTCATCGTTTTCAC 597
 DB 1256 GACCTGAGTGTGGAACGCTGATGTCATGCGGACAGGAACATTTTCCATCGCTTTGAC 1315
 QY 598 AATATCAATCAATAATACATGTCGCAAGTAGCTCAGAGAAATTTCTCAA 657
 DB 1316 AAGTTTAATGCCAAATACAAACCTATTGGGGAGTCGCTCTCCGAGAGATCTTCAACAG 1375

QY 658 CAAGATAATCTTATTCAAGGGCGTTTTCTTGCTGAGTTGACAAAGCAAGTTTTTCTCTGAC 717
 DB 1376 ACGGACAAACAGGGTATCTCTGGGAAGTACTTTGCTCACATCATCAAGGAGGTGATGTGACAG 1435
 QY 718 CTTTCTGCTAGCAAAATATCAGATGCGCAAAATATAGGATTTCAATCTACGGAAGAAACAG 777
 DB 1436 CTGGAGGAGCAAAATATCCAGATGCGAGCTGCGGCTCTCAATTTACGGGGCTCGAGG 1495
 QY 778 AGTGAATGGGACCAACTTTGCAAGTTGGATAGTGAACAAATGAATTTGCACAGTGGAAATGTT 837
 DB 1496 GATGAGTGGGACAAAGCTGGCGCTGGCCGCTCATGACCGCGTGCACCTCCCCAACGCTG 1555
 QY 838 GTCTGGCTGGTTTCAGATTTCCACGCTTATATATATGTTGTAACAAGAAATGGGTATCTTACA 897
 DB 1556 CGCTGGCTGGTGCAGGTGCGCCGCTCTTTGATGTGTACCGGTACCAAGGGCGCAGCTGGCC 1615
 QY 898 TCATTCCAAAATCTCTTGACAAACATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 957
 DB 1616 AACTTCCAGGAGATGCTGGAGAACATCTCTCTGACCTGTTGAGGGCCACTGTGCACCT 1675
 QY 958 GCTTCACACCCACAGCTCCATGTCTCTCTGAAGCAGGTTGTAGGTTGGACCTGTGTGAT 1017
 DB 1676 GCCAGCCACCCGGAAGTCACTCTCTCTTAGACGACGTGGATGTTTGGACAGCTGGAT 1735
 QY 1018 GATGAAGTAACACAGAAAGGCTCCAAACAAGC---ACATGCCCCACACCTGAAACAGTGG 1074
 DB 1736 GATGAGTCCAAAGCTGAAACCATGTCTTCAACCTGGAGAGCCCCCTGCTGAGGGCTGG 1795
 QY 1075 ACCATGTTTCAACCTGCAATTTTCAATATATGGTACTACTGTCTATCTATCTAATCTATTC 1134
 DB 1796 GTGGAGGAGGACAAACCCCTATGCTTACTACCTGTACTACCTTTTGGCAACATGGCC 1855
 QY 1135 ACCCTAAACAAAGCTGCGTGAAGTCAAGGGAATGACCACTATCAAAATTCGCTCCACATGT 1194
 DB 1856 ATGTTGAACACCTGCGCAGCAGAGGGGCTTCCACAGCTTTGTGCTCAGGCGCACCTGT 1915
 QY 1195 GGAGAGGCTGGAGATGTTGATCATCTTGGCAGCGACATTTCTTCTGTGTCACAAACATATCA 1254
 DB 1916 GGGAGGGCTGGGCCCATCCACCACCTGGTGTGAGCTTTCATGCTGGCTGAGAACATTTCC 1975
 QY 1255 CATGGAATTAATCAAGGAAGTCTCTGTGCTTCTAGTACTTGTACTATCTTGGTCAAT 1314
 DB 1976 CACGGCTCTCTTCTGCGCAAGGCCCCCTCTCTGAGTACCTGTACTCTGCGCCAGATC 2035
 QY 1315 GGTCTGGCGATGTCGCCATTCAGCAACAACTCTTATTTCTTGTACTATCATCGCAACCT 1374
 DB 2036 GGCAATCGGCATGTCTCCGCTCAGCAACAAAGAGCTTCTCTCAGCTATCAACGGAATCG 2095
 QY 1375 TTTCCAAAGCTTCTTCCAAACGAGGTCTGAATGTCTCATTTATCTACGGATGACCCCTTTGCAA 1434
 DB 2096 CTACCGGAGTACTGTCTCCCGGGCTCATGTCTCTCCCTGTCCACTGATGATCCTCTTGCAG 2155
 QY 1435 ATTCACTGCAAAAGAACCAATGTTGGGAAGAAATACAGCATTTGCTTCTGCTGTGGAG 1494
 DB 2156 TTCCACTTCCAAAGGAGCGCTGATGAGGAGTACAGCATCGCCACCCAGGTGTGGAAG 2215
 QY 1495 CTGAGTCTCTGATTTATGGAATTCGAGGAACTCTGTTTACCAATCTGGGTTTCA 1554
 DB 2216 CTGAGCTCTCGATATGTGTGAGCTGGCCCGCAACAGCGTGTCTCATGAGCGGCTTCTCG 2275
 QY 1555 CATGCTCTCAAGGGCGCACTGGATTTGGTAAGAACTACTTTCAAAAGAGGACCTGCTGGAAT 1614
 DB 2276 CACAAGGTAAAGAGCCACTGCTGGGACCCCACTATACCAAGGAAGGCCCTGAGGGGAAT 2335
 QY 1615 GATATTCAGAACCAATGTACCGCATCATAGGGTTCATTTAGAGAGATGATCTGGAGA 1674
 DB 2336 GACATCCCGCGGACCAATGTGCCAGACATCCGCGTGGGCTACCGCTACGAGACCTGTGC 2395
 QY 1675 AATGAATGAAGT 1688
 DB 2396 CAGGAGCTGGCGCT 2409

```
RESULT 13
ABL17219
ID ABL17219 standard; DNA; 2250 BP.
XX AC
XX ABL17219;
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 3130.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX Claim 1; SEQ ID NO 3130; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (ABBS7737-ABBS72072).
XX
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 2250 BP; 568 A; 622 C; 545 G; 515 T; 0 other;
XX
XX Query Match 24.1%; Score 459.4; DB 23; Length 2250;
XX Best Local Similarity 57.1%; Pred. No. 2.2e-75;
XX Matches 857; Conservative 0; Mismatches 641; Indels 3; Gaps 1;
XX
XX 152 AAGATTGTACGGAGACATTATCTCTGTGCTGATGCTCAACCTTCTTCACTGACTTC 211
XX 490 ACGAAAGCAGTGATGATCAATATCGAATATCCAGATATGAGCAATTCGTAACGACATGC 549
XX 212 ATTATATTTCTCCGAGTAACGGCTGCGAGGAACAACAAGAACTGTCTGCCATAATCGGTAA 271
XX 550 AGGTCAATGTCATATGATGCGGATGGGCCATTGAAATCTCTCTGATTCGAGCTCT 609
XX 272 ATCTCTTGTAGCATAAGTTTAAATTCATCTGATGTTAAATGCGGATAGGAAATTTCTTG 331
XX 610 GTTATCTGTCTCCAAAGTATCAGATGCACGTGCTCTCAACGAGCTCGGTGAGCTGGCTG 669
XX 332 CCAGAGACTGCCACATCGTATTTTACAAATGTCCAGAGAGTTGACATCATGTTTC 391
XX 670 CCAGAGAGCGGTGCGCATCGGATTTTACAAACACCCGCAAGGTGGACACCCATCC 729
XX 392 ATCATTCAGCATGCAATCAAAACATCTGTTGAGGTTTCATAAAATCCAAACTAAGAA 451
XX
```



```
QY 853 ATTCACGCTTATATATGTGTACAGAAATGGGTATGTTACATCATTTCCAAATCTT 912
Db |||||
QY 4573 GTTCCAGGANTCATGATGTGTTCCGTTCCAGAAATTTCTTCCACATTTTGGAAATG 4632
Db |||||
QY 913 CTTGACAAATTTCTCTCTCTTTTGGAGTTACTATTGATCCAGCTTTCACACCCACAG 972
Db |||||
QY 4633 CTGAGAAATATTTTCATGCCAGTGTGTTGAGGCCACCATCAACCCAGGCTGACCCAGAA 4692
Db |||||
QY 973 CTCCATGCTCTTCTCAAGAGGTTGTAGGGTTGGACCTGGTTGATGATGAAGTAAACCA 1032
Db |||||
QY 4693 CTCAAGTGTCTTCTCAAGCATATCACTGGCTTTGACAGTGTGATGATGATGATGATGAT 4752
Db |||||
QY 1033 GAAAGCC---GTCCACAAAGACATGCCACACCTGACAGTGGACCAATGTGTTCAAC 1089
Db |||||
QY 4753 AGTGGCCACATGTTCTCCAGAGTCCCAAGCCAGGAGTGGACATGGAAGAAT 4812
Db |||||
QY 1090 CTGCAATTTTCTATTTATGCTACTACTGTATGCTTAATTTTACCCCTAAACAGCTG 1149
Db |||||
QY 4813 CCATCTTACATTTACTATGCTACTATGATGATGATGATGATGATGATGATGATGATGAT 4872
Db |||||
QY 1150 CGTGAATCAAGGAATGACCACTATCAAAATTCCTGTCACATGCTGAGAGGCTGAGAT 1209
Db |||||
QY 4873 AGAAGGAAACGAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4932
Db |||||
QY 1210 GTTGATCACTTGGCAGGACATTTCTCTCTGTCACACATATCACATGGAATTAATCTA 1269
Db |||||
QY 4933 CTACCCATCTCATGACATTCATGATGATGATGATGATGATGATGATGATGATGATGAT 4992
Db |||||
QY 1270 AGAAGTCTCTGCTGCTCAGTACTTGTACTTCTTGGTCAATGCTGCTGCGATGCTC 1329
Db |||||
QY 4993 AAAAGAGTCCGCTGCTACAGTACTTGTGTTTCTTAGCCCAATTCCTGCGCATGCTCA 5052
Db |||||
QY 1330 CATTGAGCAAACTCTCTTATTTCTGATCATCATGCTGCAACCTTTTCAAGGTTCTTC 1389
Db |||||
QY 5053 CCACTAAGTAACAAATAGCCTTATTTCTAGATATGCAAAATTCCTTTTGGATTTCTT 5112
Db |||||
QY 1390 CAACGAGTCTGAATGTCATATTCATGATGATGATGATGATGATGATGATGATGATGAT 1449
Db |||||
QY 5113 CAGAAAGGCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5172
Db |||||
QY 1450 GAACCAATGCTGGAAGAATACAGCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1509
Db |||||
QY 5173 GAGCCCTTAATGAAGAATATGCTTATTTGCTGCAAGTCTTCAAGCTGAGCACTGTGAT 5232
Db |||||
QY 1510 TTATGGAATTCGAGGAATCTGTTTACCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1569
Db |||||
QY 5233 ATGTGCGAAGTGGCAAGAACAGTGTCTTGCAGTGTGGAATTTCTCATGAGGAGAAGTA 5292
Db |||||
QY 1570 CACTGGATTTGGTAAGAACTACTTTCAAAAGAGGACCTGCTGGAATGATATTCACGAAC 1629
Db |||||
QY 5293 AAGTTTCTGGCGACAAATTAACCTTGGAGAGGCTGCTGGAATGATATTCGGAGGACA 5352
Db |||||
QY 1630 AATGTACCGCACATCAGGGTTCAATTTAGAGAGATGATCTGGAGAAATGAATGAACATA 1689
Db |||||
QY 5353 AATGTAGCCCAATCCGATGGCTATCGCTATGAAACCTGGTGTATGAACTCAATTTA 5412
Db |||||
QY 1690 GT 1691
Db 5413 AT 5414

RESULT 15
AAS66419
ID: AAS66419 standard; cDNA; 5442 BP.
XX
AC AAS66419;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #2223.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
```

```
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
FN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
XX
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX
XX P-PSDB; ABG02232.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity -
XX
PS Claim 1; SEQ ID No 2223; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantifying a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human
XX diagnostic coding sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
```

Sequence 5442 BP; 1626 A; 1147 C; 1281 G; 1388 T; 0 other;

```
Query Match 24.0%; Score 458.8; DB 23; Length 5442;
Best Local Similarity 57.1%; Pred. No. 3e-75;
Matches 857; Conservative 0; Mismatches 642; Indels 3; Gaps 1;

QY 193 ACCTCTTCTCACTGACTTGCTATATATCTCCAGTAAACGGCTGAGGGAACAACAAGACT 252
Db |||||
QY 3913 ACCTTCTTAGACGATATGATTTTACTTGTCTTAATGCTCAAGGACCTGTTAAGACC 3972
Db |||||
QY 253 GTCTGCCAATATCGGTTAAATCTTCTTGAGCAATAAGTTTAAATTCATCTGATGTTAAAT 312
Db |||||
QY 3973 TATACCCACCGCGGCTGAAGTTCTCTCTCCAGAGTTCAGGTCATCAGATGCTTAAC 4032
Db |||||
QY 313 GCGGATAGGAATTTCTTCCCGAGAGACTGCCACATCTGATTTTACATGTCAGS 372
Db |||||
QY 4033 GAGATGGACGAGTTAAGAGAGCTGAAACAAACCCCGGAGATTTTATACCTGACG 4092
Db |||||
QY 373 AAGGTTGACACTCATGTTTCTCATCATTCAGCATCATCAATCAAAAACATCTGTTGAGGTT 432
Db |||||
QY 4093 AAGGTTGACACCCATATCCATCGAGCGCTTGATGAGACCAACAACATCTGCTGCTTT 4152
Db |||||
QY 433 ATAAATCCAACTAAGAAAAAGAACTGATGAGGTGCTCATTTTCAGAGATGCTATTAT 492
Db |||||
```

Db 4153 ATTAGAATCTTACCAATTGATGCTGACAGAGTGGTCTATAGACCAAAAGAGAGAT 4212
Qy 493 ATGACTTTAAAGAGGTTTTGAGAGCTTGGACTTAACCTGGGTATGATCTGTAATGTGAT 552
Db 4213 CTGACCTTAAGGAACTTTTGTCTAAATTAATAATGATCATCTTATGACCTGATGTTGAT 4272
Qy 553 TTGCTAGATGCCATGACAGCAAAAGCAGATTTTCATCGTTTTTGACAAATTCATCTAA 612
Db 4273 TCTTGGATGTTTCATGCTGGACGCCAGACCTTCCAGCGTTTTGATAAGTTCAATGACAA 4332
Qy 613 TACAATCCATGTGGCCAAAGTAGGCTCAGAGAAATTTCTCAAAACAGATAATCTTATT 672
Db 4333 TATATCTGTAGGAGCAGTACGGACCTCTACTTGAAGACAGACATTTACATT 4392
Qy 673 CAAAGCCGTTTTCTGCTGAGTGAACAAAGCAAGTTTTCTTGACCTTTCTCTAGCAAA 732
Db 4393 AATGGGAATAATTTGGCCACTATCATCAAGGAGGTAGTGGCGACCTGGTGGAGCCAAG 4452
Qy 733 TATCAGATGCCAGATATAGGATTTCAATCTACGGAGGAAACAGATGAATGGAGCCAA 792
Db 4453 TACCAGCATGCTGAGCCCGCCCTGTCTCATATAGCCGCACTCTGATGAGTGGAGCAAA 4512
Qy 793 CTTCGAAGTTGATAGTGAACAAATGAATGCAAGTGGAAATGTTGCTGGCTGGTTGAG 852
Db 4513 CTCTCTCTGTTGCTGTGCAATCGCATCCATGCCCCAATGACATGGATGATCCAG 4572
Qy 853 ATTCACCGCTTATATAATGTGTAAGGAAATGGGTATCGTTACATCATTTCCAAAATCTT 912
Db 4573 GTTCCAGGATCTATGATGTGTTCCGTTCCAAAGAAATTTCTTCCACATTTTGGAAAATG 4632
Qy 913 CTGTGACAAATTTTGGTCTCTTTTGGAGTTACTATTTGATCCAGTTCACACCCACAG 972
Db 4633 CTGGAGAAATTTTTCATGCGAGTGTGTGAGGCCACCATCAACCCCGAGCTGACCCAGAA 4692
Qy 973 CTCCATGCTCTCTGAAGCAGGTTGTAGGTTGGACCTGGTTGATGAGAAAGTAAACCA 1032
Db 4693 CTCAGTGTCTCTCAAGCATATCACTGSCTTTGAAGTGGATGATGATGCCAAACAC 4752
Qy 1033 GAAAGC---GTCCAAAGACACATGCCACACCTGAAACAGTGGACCAATGTGTTCAAC 1089
Db 4753 AGTGGCCACATGTTCTCTCCAGAGTCCCAAGCCCCAGAGTGGACATTTGAAAAGAT 4812
Qy 1090 CTGTGATTTTCAATATGCGTACTACTGCTATGCTATTAATTCACCTAAACAGCTG 1149
Db 4813 CCATCTTACACTTACTATGCTTACTACATGTATGCAAAACATCATGCTGTCAACAGCCTG 4872
Qy 1150 CGTGAGTCAAGGGAATGACCACTATCAAAATTCGTCACATGCTGGAGAGCTGGAGAT 1209
Db 4873 AGAAAGGAACGAGGATGAATACGTTTCTGTTCCGACCTCACTGTGGAGAGCTGGAGCC 4932
Qy 1210 GTTGATCACTTGGCAGCGACATTTCTTCTGTCAACATATCACATGGAATTAATCTA 1269
Db 4933 CTCACCCATCTCATGACAGATTCATGATAGCAGATGATATCTCTCATGGCCTAAATTTA 4992
Qy 1270 AGGAAGTCTCTGTGCTCAGTACTTGTACTATCTTGGTCAATTTGGTCTGGCGATGCC 1329
Db 4993 AAAAAAGTCCCGTGTACAGTACTTGTGTTTCTTAGCCCAATTTCCCATGCCATGTCA 5052
Qy 1330 CCATTGAGCAACAATCTCTTATTTCTGACTATCATCGCAACCCCTTTTCCAGCTTCTTC 1389
Db 5053 CCACCTAAGTAACTAGGCTATTTCTAGAGTATGCCAAAATCTCTTTTGGATTTCTCT 5112
Qy 1390 CAAAGAGTCTGAATGTCTCATTATCTACGATGACCCCTTTTGCAAAATTCACCTGACAAA 1449
Db 5113 CAGAAAGGCTAATGATCTCACTGTCTACAGATGACCAATGCAATTCACCTTTACCAAG 5172
Qy 1450 GAAACATTTGGTGAAGAAATACAGCATTTGCTGCTTGGTGGAGCTCAGTTCTTTGAT 1509
Db 5173 GAGCCCTTAATGGAAGAAATATGCTATTTGCTGCAAGTCTTCAAGCTGAGCACCTGTGAT 5232
Qy 1510 TTATGGAAATTTGGAGGAGTCTCTTTTACCAATCTGGGTTTTTCAATGCTCTCAAGCG 1569
Db 5233 ATGTGGAAGTGGCAAGAACAGTGTCTGTGAGTGTGAAATTTCTCATGAGGAGAAAGTA 5292

Qy 1570 CACTGATTTGTAAGAACTACTTTCAAAAGAGGACCTGCTGAAAATGATATTTCAAGAAC 1629
Db 5293 AAGTTTCTGGCGCAAAATTACTTTGAGGAAGGCCCTGCTGAAAATGATATCCGGAGGACA 5352
Qy 1630 AATGTACCGCACATCAGGTTTCAATTTAGAGAGATGATCTGGAGAAATGAAATGAACTA 1689
Db 5353 AATGTAGCCCAATCCGCATGCGCTATGCGCTATGAAAACCTGGTGTATGAACTCAATTTA 5412
Qy 1690 GT 1691
Db 5413 AT 5414

Search completed: December 15, 2003, 18:10:04
Job time : 525 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	79.6	4.2	1248	4	US-09-489-847-101	Sequence 101, Appl
2	77.2	4.0	1701	4	US-09-996-243-114	Sequence 114, Appl
3	76.2	4.0	2806	4	US-09-653-839-9	Sequence 9, Appl
4	74.6	3.9	1641	1	US-08-300-903A-8	Sequence 8, Appl
5	74.6	3.9	1641	4	US-08-988-197-8	Sequence 8, Appl
6	74.4	3.9	1927	4	US-09-336-536-66	Sequence 66, Appl
7	73.6	3.9	2550	6	5258287-23	Patent No. 5258287
8	73.4	3.8	664	4	US-09-904-615-66	Sequence 66, Appl
9	73.4	3.8	1864	4	US-09-149-476-130	Sequence 130, Appl
10	73.2	3.8	1674	4	US-09-996-243-300	Sequence 300, Appl
11	72.8	3.8	1098	3	US-09-248-335-35	Sequence 35, Appl
12	72.6	3.8	1114	4	US-09-152-060-41	Sequence 41, Appl
13	72.4	3.8	1342	4	US-09-489-847-89	Sequence 89, Appl
14	72.2	3.8	2447	2	US-09-014-963-14	Sequence 14, Appl
15	72	3.8	1882	3	US-09-370-253-1	Sequence 1, Appl
16	71.6	3.7	1069	4	US-09-372-422A-7	Sequence 7, Appl
17	71.4	3.7	441	4	US-09-601-537-10	Sequence 10, Appl
18	71.4	3.7	4121	4	US-09-601-537-9	Sequence 9, Appl
19	71	3.7	331	4	US-09-482-273-31	Sequence 31, Appl
20	71	3.7	3715	4	US-09-234-245-1	Sequence 1, Appl
21	70.8	3.7	578	4	US-09-602-877A-95	Sequence 95, Appl
22	70.8	3.7	1602	1	US-08-530-950-3	Sequence 3, Appl
23	70.8	3.7	1602	3	US-08-888-429A-3	Sequence 3, Appl
24	70.8	3.7	1602	3	US-09-149-873-3	Sequence 3, Appl
25	70.8	3.7	1602	4	US-09-057-009-3	Sequence 3, Appl
26	70.8	3.7	3238	3	US-08-123-934A-5	Sequence 5, Appl
27	70.8	3.7	3328	5	PCT-US94-10080-5	Sequence 5, Appl

Patent No. 6478825
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2730PIC13
CURRENT APPLICATION NUMBER: US/09/996,243
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429

;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090431
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090435
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090576
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Query Match 4.0%; Score 77.2; DB 4; Length 1701;
Best Local Similarity 72.5%; Pred. No. 6.5e-09;
Matches 100; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 1773 TGTACTGCTATGGCAATTTATCTTCATGTTTGGTATGCTTCTCTATCTATGCAAT 1832
Db 1554 TTTTGCTAATTTGTTACTTTTCTTGCTAATTTGGAAGATTAATCTATTTTATAA 1613

QY 1833 TCACTTCGAATCTCAAAAAA 1892
Db 1614 TTATGCTAAGATTAAAAA 1673

QY 1893 AAAAAA 1910
Db 1674 AAAAAA 1691

RESULT 3
US-09-653-839-9

;; Sequence 9, Application US/09653839
;; Patent No. 6433153
;; GENERAL INFORMATION:
;; APPLICANT: Donoho, Gregory
;; APPLICANT: Turner, C. Alexander Jr.
;; APPLICANT: Nehls, Michael
;; APPLICANT: Friedrich, Glenn
;; APPLICANT: Zambrowicz, Brian
;; APPLICANT: Sands, Arthur T.
;; TITLE OF INVENTION: No. 6433153el Human Calcium Dependent Proteases
;; TITLE OF INVENTION: and Polynucleotides Encoding the Same
;; FILE REFERENCE: LEX-0038-USA
;; CURRENT APPLICATION NUMBER: US/09/653.839
;; CURRENT FILING DATE: 2000-09-01
;; PRIOR APPLICATION NUMBER: US 60/152,057
;; PRIOR FILING DATE: 1999-09-02
;; NUMBER OF SEQ ID NOS: 9
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 9
;; LENGTH: 2806
;; TYPE: DNA
;; ORGANISM: homo sapiens
US-09-653-839-9

Query Match 4.0%; Score 76.2; DB 4; Length 2806;
Best Local Similarity 76.9%; Pred. No. 1.3e-08;
Matches 93; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 1790 TTATACITTCATGTTTGGTATGCTTCTCTATCTATGCAATTCGAATTCGAATTCGA 1849
Db 2661 TTTTCTCCCATCTGTGATACATTCTAATAATAGCACATGCCATTCGCAAAAAA 2720

QY 1850 AAAAAA 1909
Db 2721 AAAAAA 2780

QY 1910 A 1910
Db 2781 A 2781

RESULT 4
US-08-300-903A-8
;; Sequence 8, Application US/08300903A
;; Patent No. 5591630
;; GENERAL INFORMATION:
;; APPLICANT: Anderson, Dirk M
;; APPLICANT: Gird, Judith G
;; TITLE OF INVENTION: Interleukin-15 Receptors
;; NUMBER OF SEQUENCES: 15
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Immunex Corporation
;; STREET: 51 University Street
;; CITY: Seattle
;; STATE: Washington
;; COUNTRY: USA
;; ZIP: 98101
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: Apple Macintosh
;; OPERATING SYSTEM: Apple Operating System 7.1
;; SOFTWARE: Microsoft Word for Apple, Version 5.1a
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/300.903A
;; FILING DATE: 06-SEPTEMBER-1994
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: USSN 08/236,919
;; FILING DATE: 06-MAY-1994
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Perkins, Patricia Anne
;; REGISTRATION NUMBER: 34,695

```
; REFERENCE/DOCKET NUMBER: 2822-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-587-0430
; TELEFAX: 206-233-0644
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1641 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..839
; US-08-300-903A-8

Query Match 3.9%; Score 74.6; DB 1; Length 1641;
Best Local Similarity 85.6%; Pred. No. 2.6e-08;
Matches 83; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1814 TTCCTTATCTATGCGCAAAATTCACACTTCGAACTTCACAAAAA 1873
Db 1516 TACCTTGATACAAATAAACAATCTATTTTCAATAAAAAA 1575

QY 1874 AAAAAA 1910
Db 1576 AAAAAA 1612

RESULT 5
US-08-988-197-8
; Sequence 8, Application US/08988197
; Patent No. 6548065
; GENERAL INFORMATION:
; APPLICANT: Anderson, Dirk M
; APPLICANT: Giri, Judith G
; TITLE OF INVENTION: Interleukin-15 Receptors
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Apple Operating System 7.1
; SOFTWARE: Microsoft Word for Apple, Version 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/988,197
; FILING DATE:
; CLASSIFICATION: 121097
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/300,903
; FILING DATE: 06-SEPTEMBER-1994
; APPLICATION NUMBER: USSN 08/236,919
; FILING DATE: 06-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,695
; REFERENCE/DOCKET NUMBER: 2822-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-587-0430
; TELEFAX: 206-233-0644
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1641 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

; REFERENCE/DOCKET NUMBER: 2822-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-587-0430
; TELEFAX: 206-233-0644
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1641 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..839
; US-08-988-197-8

Query Match 3.9%; Score 74.6; DB 4; Length 1641;
Best Local Similarity 85.6%; Pred. No. 2.6e-08;
Matches 83; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1814 TTCCTTATCTATGCGCAAAATTCACACTTCGAACTTCACAAAAA 1873
Db 1516 TACCTTGATACAAATAAACAATCTATTTTCAATAAAAAA 1575

QY 1874 AAAAAA 1910
Db 1576 AAAAAA 1612

RESULT 6
US-09-336-536-66
; Sequence 66, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
; APPLICANT: Leiby, K.
; APPLICANT: McKay, C.
; APPLICANT: Bossone, S.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-144
; CURRENT APPLICATION NUMBER: US/09/336,536
; CURRENT FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 66
; LENGTH: 1927
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified base
; LOCATION: all "n" positions
; OTHER INFORMATION: n=a, c, g, or t
; US-09-336-536-66

Query Match 3.9%; Score 74.4; DB 4; Length 1927;
Best Local Similarity 74.4%; Pred. No. 3e-08;
Matches 93; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 1786 GGAATTATATCTTCATGTTTGGTATGCTTTCTTATCTATGCAAAATTCACACTTCGAACT 1845
Db 1733 GGACCTTCTCCACATGTTTGTATGCAACATTTTGCATTAAAGGAATCCANAAA 1792

QY 1846 TCANAAAAA 1905
Db 1793 AAAAAA 1852

QY 1906 AAAAA 1910
Db 1853 AAAAA 1857

RESULT 7
5258287-23
; Patent No. 5258287
; APPLICANT: BAXTER, ROBERT C.; WOOD, WILLIAM I.
; TITLE OF INVENTION: DNA ENCODING AND METHODS OF PRODUCTION
; OF INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN BP53
; NUMBER OF SEQUENCES: 58
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/171,623
; FILING DATE: 22-MAR-1988
```


; Patent No. 6478825
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Auscin L.
; APPLICANT: Kijavini, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas P.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730PIC13.
; CURRENT APPLICATION NUMBER: US/09/996,243
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429

; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 3.8%; Score 73.2; DB 4; Length 1674;
Best Local Similarity 71.6%; Pred. No. 5.5e-08;
Matches 96; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 1777 GCTGCTATGGAAATTAATCTCAATGTTTGGTATGCTTTCCTTATCTATGCGCAAAATCAA 1836
Db |||||
1514 GCTTCTCAGGACGAGGACATCTTGGCAGTCTGGCTTGGCCATTAAATGGAACCTGAA 1573
QY |||||
1837 CTTTCGAACTTCAAAAAA 1910
Db |||||
1574 GGCACAAAAA 1647
QY |||||
1634 AAAAAA 1647

RESULT 11
US-09-248-335-35

; Sequence 35, Application US/09248335
; Patent No. 6096504
; GENERAL INFORMATION:
; APPLICANT: MCGONIGLE, BRIAN
; APPLICANT: O'KEEF, DANIEL
; TITLE OF INVENTION: PLANT GLUTATHIONE-S-TRANSFERASE ENZYMES
; FILE REFERENCE: CL-1128-A
; CURRENT APPLICATION NUMBER: US/09/248,335
; CURRENT FILING DATE: 1999-02-10
; EARLIER APPLICATION NUMBER: 08/924,759
; EARLIER FILING DATE: 1997-September-05
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 35
; LENGTH: 1098
; TYPE: DNA
; ORGANISM: maize
US-09-248-335-35

Query Match 3.8%; Score 72.8; DB 3; Length 1098;
Best Local Similarity 79.6%; Pred. No. 6e-08;
Matches 86; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 1803 TTTGCTATGCTTTCCTTATCTATGCGCAAAATCAAATTCGAACTTCAAAAAA 1862
Db |||||
965 TCTTTTATTTGTTGCTTCTTTAAACAAGAGTTATATTTTACCATCAAAAAA 1024
QY 1863 AAAAAA 1910
Db |||||
1025 AAAAAA 1072

RESULT 12:

; Sequence 41, Application US/09152060
; Patent No. 6448230
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: PZ003PI.US
; CURRENT APPLICATION NUMBER: US/09/152,060
; CURRENT FILING DATE: 1998-09-11
; EARLIER APPLICATION NUMBER: PCT/US98/04858
; EARLIER FILING DATE: 1998-03-12
; EARLIER APPLICATION NUMBER: 60/040,762
; EARLIER FILING DATE: 1997-03-14
; EARLIER APPLICATION NUMBER: 60/040,710
; EARLIER FILING DATE: 1997-03-14
; EARLIER APPLICATION NUMBER: 60/050,934
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,100
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,357
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,189
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/057,765
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/068,368
; EARLIER FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
; LENGTH: 1114
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-152-060-41

Query Match 3.8%; Score 72.6; DB 4; Length 1114;
Best Local Similarity 73.2%; Pred. No. 6.7e-08;
Matches 90; Conservative 2; Mismatches 31; Indels 0; Gaps 0;

Search completed: December 15, 2003, 21:08:09
Job time : 142 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2003, 20:04:30 ; Search time 605 Seconds
(without alignments)
10492.685 Million cell updates/sec

Title: US-10-019-633-1
Perfect score: 1910
Sequence: 1 ccacgcgcgggtgtctcc.....aaaaaaaaaaaaaaaaaaaa 1910

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2201672 seqs, 1661799599 residues

Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1028	53.8	1803	10	US-09-938-842A-1715
2	471.6	24.7	3386	13	US-09-873-319-446
3	471.6	24.7	3386	13	US-09-960-706-706
4	458.8	24.0	5442	15	US-10-119-926-83
5	422	22.1	3915	13	US-10-205-219-188
6	282	14.8	282	9	US-09-294-0938-2101
7	177.6	9.3	278	9	US-09-923-876-5583
8	145.2	7.6	500	11	US-09-991-936-72
9	87.2	4.6	438	10	US-09-930-352-3507
10	77.2	4.0	255	13	US-09-960-213-630
11	77.2	4.0	1672	13	US-10-063-735-17
12	77.2	4.0	1672	13	US-10-063-526-17
13	77.2	4.0	1672	13	US-10-063-586-17
14	77.2	4.0	1672	13	US-10-063-510-17
15	77.2	4.0	1672	13	US-10-063-514-17

16	77.2	4.0	1672	13	US-10-063-516-17	Sequence 17, Appl
17	77.2	4.0	1672	13	US-10-063-523-17	Sequence 17, Appl
18	77.2	4.0	1672	13	US-10-063-527-17	Sequence 17, Appl
19	77.2	4.0	1672	13	US-10-063-528-17	Sequence 17, Appl
20	77.2	4.0	1672	13	US-10-063-529-17	Sequence 17, Appl
21	77.2	4.0	1672	13	US-10-063-536-17	Sequence 17, Appl
22	77.2	4.0	1672	13	US-10-063-540-17	Sequence 17, Appl
23	77.2	4.0	1672	13	US-10-063-546-17	Sequence 17, Appl
24	77.2	4.0	1672	13	US-10-063-562-17	Sequence 17, Appl
25	77.2	4.0	1672	13	US-10-063-564-17	Sequence 17, Appl
26	77.2	4.0	1672	13	US-10-063-565-17	Sequence 17, Appl
27	77.2	4.0	1672	13	US-10-063-568-17	Sequence 17, Appl
28	77.2	4.0	1672	13	US-10-063-570-17	Sequence 17, Appl
29	77.2	4.0	1672	13	US-10-063-577-17	Sequence 17, Appl
30	77.2	4.0	1672	13	US-10-063-579-17	Sequence 17, Appl
31	77.2	4.0	1672	13	US-10-063-581-17	Sequence 17, Appl
32	77.2	4.0	1672	13	US-10-063-582-17	Sequence 17, Appl
33	77.2	4.0	1672	13	US-10-063-583-17	Sequence 17, Appl
34	77.2	4.0	1672	13	US-10-063-584-17	Sequence 17, Appl
35	77.2	4.0	1672	13	US-10-063-587-17	Sequence 17, Appl
36	77.2	4.0	1672	13	US-10-063-589-17	Sequence 17, Appl
37	77.2	4.0	1672	13	US-10-063-591-17	Sequence 17, Appl
38	77.2	4.0	1672	13	US-10-063-592-17	Sequence 17, Appl
39	77.2	4.0	1672	13	US-10-063-593-17	Sequence 17, Appl
40	77.2	4.0	1672	13	US-10-063-596-17	Sequence 17, Appl
41	77.2	4.0	1672	13	US-10-063-597-17	Sequence 17, Appl
42	77.2	4.0	1672	13	US-10-063-600-17	Sequence 17, Appl
43	77.2	4.0	1672	13	US-10-063-602-17	Sequence 17, Appl
44	77.2	4.0	1672	13	US-10-063-604-17	Sequence 17, Appl
45	77.2	4.0	1672	13	US-10-063-606-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1
US-09-938-842A-1715
; Sequence 1715, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRI1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1715
; LENGTH: 1803
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1715

Query Match 53.8%; Score 1028; DB 10; Length 1803;
Best Local Similarity 76.2%; Pred. No. 3.6e-231;
Matches 1265; Conservative 0; Mismatches 395; Indels 0; Gaps 0;
QY 13 GTTGCTCCATGGGAGAGGAGTCAATATGACCCCTGACTCTCCAAAAACCTAACCCCAAC 72
Db 100 GTTGCAACATGGGAAAAAGAGTCAATCTGCTAGTCTCTCCAAAGCTTAATACAGAG 159
QY 73 CGGTTCATTATGCTGCTGAACCAAGTCAGAGCATGTTTCCAAACTGTTGATGGCGTT 132
Db 160 CCATTTGCACACTATCTCTCAGGAAAAATCTGATCATCTGTTTGAGATCAAGATGGGTT 219

```
QY 133 ATCCATGTTTATGCGGATAAAGATTGTACGGAGACATTTATCTCTGTGCTGATGCTACA 192
Db 220 GTCCACGCTGTTGCAATTAAGAGATGCAAAAGAGATCTCTTCCCGTAGCTGATGCCACA 279
QY 193 ACCTCTTCACTGACTTGCAATTATATTTCTCCGAGTAAACGGCTGCGAGGAAACACAAAGACT 252
Db 280 GCGTTTTTCACTGACTTGCACTCACTGACTCAAGTCAATGCTGCGAGGAAACATCCGACT 339
QY 253 GTCTGCATAATCGGTTAAATCTTCTTGAGCATAGTTTAAATTCATCTGATGTTAAAT 312
Db 340 TTGTGACCGCTGAGTGTCTTCTAGAACAGAAATTTAATCTCCATTTGATGCTTAAT 399
QY 313 GCGGATAGGGAATTTCTTCCGAGAAGACTGCCCCACATCGTGATTTTACAAATGTCAGG 372
Db 400 GCGGATTAAGAAATTTCTGCTCAAAAAGTGACCCATCGTGATTTTATACGTTAGG 459
QY 373 AAGTTTGACACTCATGTTTCATCATTCAGCATGATGAATCAAAAACATCTGTTGAGGTTT 432
Db 460 AAGTCGACACTCATGTCATCATTCAGCTTGCAATGAACACGAAACACCTTTTAAGGTTT 519
QY 433 ATAAATCAAACTAAGAAAGAACTGATGAGTGGTCACTTAACCTGGGTATGATCTGAATGTTGAT 492
Db 520 ATTAAGTCAAGCTCCGGAAGAAACCCGATGAGGTTGTAATATTCGAGATGGAACATAT 579
QY 493 ATGACTTTAAAGAGAGTTTGTGAGAGCTTGGACTTAACCTGGGTATGATCTGAATGTTGAT 552
Db 580 TTGACCTTGAGAGAGTTTGTGAGAGCTTGGATCTGACTGGATATGACCTGAAAGTCGAC 639
QY 553 TTGCTAGATGTCCATGCAGACAAAAGACATTTTCATCGTTTGTGACAAATTCATCTAAAA 612
Db 640 CTTTTGGATGTTTCATGCAGACAAAAGTACCTTTTCATCGTTTGTGATAAGTTCAACCTAAAG 699
QY 613 TACAATCCATGTTGGCCAAAGTAGGCTCAGAGAAATTTTCTCTCAAAACAGATATCTTATT 672
Db 700 TATAACCTTGTGTGTCAAAGTAGGCTTAGGGAATTTTCTTAAACAGGATAATCTCATC 759
QY 673 CAAGGCGGTTTCTCTGAGTTGACAAAGCAAGTTTCTCTGACCTTTCTGCTAGCAAA 732
Db 760 CAAGTCTGATTTCTTGAGATGATCAAAAGCAAGTCTTCTGACCTTGAAGCTAGTAAA 819
QY 733 TATCAGATGCGAATATATAGGATTTCAATCTACGGAAGGAAACAGAGTGAATGGGACCAA 792
Db 820 TATCAGATGCTGAATACAGATATCTATATATGCGAGAAATAGAGAGTGGGACCAA 879
QY 793 CTTGCAAGTTGATGATGAATGAATGACAGTGGAAATGTTGTTCTGGCTGGTTTCAAG 852
Db 880 CTGCTAGTTGGATTTGAAACATGATCTATACAGTGAAGATGTTGTTCTGGTTAAATTCAG 939
QY 853 ATTCCACGCTTATATAATGTGTACAAGGAAATGGGTATCGTTACATCATTTCCAAATCTT 912
Db 940 CTCCACGCTTGTACAAATTTACAAGACATGGGTAATTTGACATCGTTTCAGATATC 999
QY 913 CTTGACAAATTTGCTCTCTCTTTTGTAGGTTTACTATTGATCCAGCTTTCACACCCACAG 972
Db 1000 CTGGACAATATATTCATCTCTGTTTGAAGCCACGCTAGATCTGATTTCCCATCTCTCAG 1059
QY 973 CTCCATGCTTCTGAGCAGGTTGTAGGTTGGACCTGGTTGATGATGAATGAAGTAAACCA 1032
Db 1060 CTCCATGTTTTTTTGAAGCAGGTTGTGGATTTGATTTGATTTGATGATGAAGAACCAACCT 1119
QY 1033 GAAAGCGTCCAAACAGACATGCCACACCTGAAACAGTGGACCAATGTTTCAACCTT 1092
Db 1120 GAAAGAGCTCCCAAAACACATGCCCACTCCAGCTCAATGGAATCAACCATTCATCTCT 1179
QY 1093 GCATTTTCAATATATGCTACTACTGCTATGCTAACTTATTCACCTTAAACAGCTGCGT 1152
Db 1180 GCATTTTGTATTTATGCTACTATTGTTATGCTTAACTCTATGTTGTTAAATTAAGCTTGA 1239
QY 1153 GAGTCAAGGGAATGACCACTATCAAAATTCGCTTCACTGCTGGAGAGGCTGGAGATGTT 1212
Db 1240 GAGTCAAGGGCATGACTACTATACGCTACGACCACTTCTGGAGAGGCTGGTGACATT 1299
```

```
QY 1213 GATCATTGGCAGCGACATTTCTCTGTGTCAACAATATCAGATGGAATTAATTAAGG 1272
Db 1300 GACCACTTGGCTGCTACGTTTCTTAACATGCCATAGCATCGCACATGGAATCAATCTGCGA 1359
QY 1273 AAGTCTCTGTGCTTCACTGATCTTGTATCTTGTGTAGATTTGTTCTGGCGATGTCGCCA 1332
Db 1360 AAGTCTCTGTGCTTCACTGATCTGATCTGCTGCGCCAGATTTGGTCTGCGCATGTCAACA 1419
QY 1333 TTGACACAACTCTTATTTCTTGTGATCATCATCGCAACCTTTTCCAAAGCTTCTTCCAA 1392
Db 1420 CTGAGCAACAACTCTTGTGTTAGATTAACACCGGAACCCGTTTCTCTGTGTTTCTTTA 1479
QY 1393 CGAGTCTGAATGTCTCAATTAATCTACGATGACCCCTTTGCAAAATTCACCTGACAAAAGAA 1452
Db 1480 AGAGTCTCAATGTTTCTCTGTCTACTGATGACCCCTTCAAGATTCATTTAACTAAAGAA 1539
QY 1453 CAATTGGTGAAGAAATACAGCATTTGCTGCTGCTGTTGGAAGCTCAGTTCTTGTGATTTA 1512
Db 1540 CTTCTCGTGAAGAGTATAGCATAGCTGCATCAGTTTGGAAAGCTGAGTGCCTGTGACCTG 1599
QY 1513 TCGGAAATTCGAGGAACTCTGTTTACCAATCTGGGTTTTCACATGCTCTCAAGCGGCAC 1572
Db 1600 TCGGAGATAGCTCGTAATCTAGTGTACAGTACAGGTTTCTCACAGCCCTGAACTGCGAC 1659
QY 1573 TGGATTGTGAAGAACTACTTCAAAAGAGGACCTGCTGGAATGATTAATTCACAGAACCAAT 1632
Db 1660 TGGATTGGAAGATTTACTACAAAGAGGACCTGATGGAACGACATTCACAAACAAAC 1719
QY 1633 GTACCGCACATCAGGTTTCAATTTAGAGAGATGATCTGGA 1672
Db 1720 GTGCCACACATAAGGGTGGAGTTCGTGACACGGTATGGA 1759
```

RESULT 2

```
US-09-873-319-446
; Sequence 446, Application US/09873319A
; Publication No. US20030134324A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William E.
; APPLICANT: Kulkarni, Prakash
; APPLICANT: Getzenberg, Robert H.
; APPLICANT: Waga, Iwao
; APPLICANT: Yamamoto, Jun
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic
; TITLE OF INVENTION: Hyperplasia Using Gene Expression Profiles
; FILE REFERENCE: 44921-5029-US
; CURRENT APPLICATION NUMBER: US/09/873,319A
; EARLIER FILING DATE: 2001-06-05
; EARLIER APPLICATION NUMBER: US 60/223,323
; NUMBER OF SEQ ID NOS: 755
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 446
; LENGTH: 3386
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134324A1 M91029
US-09-873-319-446
```

```
Query Match 24.7%; Score 471.6; DB 13; Length 3386;
Best Local Similarity 58.3%; Pred. No. 2.9e-100;
Matches 847; Conservative 0; Mismatches 604; Indels 3; Gaps 1;
QY 238 GGGAAACAAGAACTGTCTGCCATAATCGGTTAAATCTTTTGAGCATAAGTTTAAATTC 297
Db 956 GGGCCCAATAAAGTCAATCTGCTACCGCGGCTGCGAGTACCTCAGCTCCAAGTTCAGATG 1015
QY 298 CATCTGATGTTAAATCGGATAGGGAATTTCTTGCCCAAGAGACTGCCCCACATCGTAT 357
Db 1016 CATGTGCTACTCAATGAGATGAAGAGCTGGCCGCCCAAGAAAGTGCCACCCGAGAT 1075
QY 358 TTTTACAATGTGAGGAAGGTTGACACTCATGTTTCAATTCAGCATGATGAATCAAAA 417
```


Db 1076 TTCTACAAATCCGAGAGTGGACACCAATCCATGCTCTGCTCGATGAACCAAG 1135
Qy 418 CATCTGTGAGTTCATAAAATCAAACTAAGAAAAGAACCTGATGAGGTGGTCAATTTTC 477
Db 1136 CATCTGTGCTTTCATCAAGCGGCAATGAAGCGCACCTGGAGGAGATCGTGACGTC 1195
Qy 478 AGAGATGGTACTTATATGACTTTAAAGAGGTTTTGAAGAGCTTGAACCTTAACCTGGGTAT 537
Db 1196 GAGCAGGGCGGTGAACAGACGCTCGGGAGGTCTTTGAGAGCATGAATCTCAAGCGCTAC 1255
Qy 538 GATCTGAATCTTATGTTGCTAGATGCTCCATGACAGACAAAGACATTTTCATGTTTGGAC 597
Db 1256 GACCTGAGTGTGACACGCTGGAATGTCATCGGACAGGAACATTTTCATGCTTTGAC 1315
Qy 598 AAATCAATCAAAATACAAATCATGTCGCAAGTAGGCTCAGAGAAATTTTCTCAAA 657
Db 1316 AAGTTTAATGCCAAATACAAACCTATTGGGAGTCCGCTCTCGAGAGATCTTCATCAAG 1375
Qy 658 CAAGATAATCTTATTCAAGCGCGTTTCTTGCTGAGTTGACAAAGCAAGTTTCTCTGAC 717
Db 1376 ACGGACAAACAGGATATCTGGGAATGATTTTGCTCACATCATCAAGGAGGTGATGTCAGAC 1435
Qy 718 CTTTCTGTAGCAAAATACAGATGCGCAGATATAGGATTTCAATCTACGGAAGAACAG 777
Db 1436 CTGGAGAGCAAAATACGAAATGAGAGCTCGGCTCTCCATTTACGGGCGCTCGAGG 1495
Qy 778 AGTGAATGGACCAACTTGAAGTTGGATGAGCAAAATGAATTCACAGTGAATGTT 837
Db 1496 GATGAGTGGACAACTGGCGCGCTGGCGCTCATGCAAGCGCTGACTCCGCCAACGTCG 1555
Qy 838 GTCTGCTGTTGAGATTCACGCTTATATGATGTCACAGAAATGGGTATCGTTACA 897
Db 1556 CGTGTGCTGAGTGGCGGCGCTCTTTGATGTCAGTACCAAGGCGCAGCTGGCC 1615
Qy 898 TCATTCCAAAATCTTCTGACACATTTTGGTCTCTTTTGGAGTTACTATTGATCA 957
Db 1616 AACTTCAGAGATGCTGGAGAACATCTTCTGCGCACTGTTGAGGCGCACTGTGCAACCT 1675
Qy 958 GCTTCAACCCACAGCTCATGTCCTTCTGAGACAGGTTGTAGGGTTGACCTGGTTGAT 1017
Db 1676 GCCAGCACCGGAATCGATCTCTTTAGAGCACGTGGATGTTTTCAGACGTGGAT 1735
Qy 1018 GATGAAGTAAACAGAAAGGCGTCCAAAGC---ACATGCCACACCTGAACAGTGG 1074
Db 1736 GATGAGTCCAAAGCTGAAACCAATGCTTCAACCTGGAGAGCCCTGCTGAGGCGTGG 1795
Qy 1075 ACCAATGCTTCAACCTGCAATTTTCAATATGCTACTACTGCTATGCTACTACTTATTC 1134
Db 1796 GTGGAGGAGAACCAACCTATGCTTACTACTACTACTACTACTACTACTACTACTACT 1855
Qy 1135 ACCCTAAACAAAGCTGCTGAGTCAAGGGAATGACCACTATCAAAATTCGCTCCACATGCT 1194
Db 1856 ATGTTGAACCACTGCGAGGAGAGGGCTCCACAGCTTTGCTGAGGCGCACTGTT 1915
Qy 1195 GAGAGGCTGGAGATTTGATCATCTTGGAGGACATTTCTTCTGTCACACATATCA 1254
Db 1916 GGGAGGCTGGGCCATCAACCACTGGTGTGAGCTTCAATGCTGGCTGAGAACAATTC 1975
Qy 1255 CATGGAATTAATCAAGGAAGTCTCTGCTGCTTCACTACTGCTACTACTACTACTACT 1314
Db 1976 CAGGGCTCTTCTGGCAAGGCCCGCTGCTGCACTACTACTACTACTACTACTACTACT 2035
Qy 1315 GGTCTGGGATGTCGCCATTTGAGCAACAACTCTCTATTTCTTGTACTATCATCGCAACCT 1374
Db 2036 GGCATGCCATGTCCTCGCTCAGCAACAAAGCTCTTCTCAGCTATCACCGGAATCG 2095
Qy 1375 TTTTCAACGTTCTTCAACAGAGTCTGAATGTCTCATTTATCTACGAGTACCCCTTTGCAA 1434
Db 2096 CTACCGGAGTACTGTCGCGGCGCTCATGGTCTCCCTGCTCCACTGATGATCCCTTGCAG 2155
Qy 1435 ATTCACTCACAAGAACCATTTGTTGGAAGATACAGCAATTTGCTGCTTCTGCTGTGGAG 1494

Db 2156 TTCCACTTCCACCAAGAGCCGCTGATGAGGAGTACAGATCGCCACCCAGGTGTGGAG 2215
Qy 1495 CTCAGTCTTGTGATTTATGCGAAATTCGAGAACTCTGTTTACCAATCTGGGTTTTCA 1554
Db 2216 CTCAGCTCTCGATATGTTGAGCTGGCCGACACAGCGTCTCATGAGCGCTTCTCG 2275
Qy 1555 CATGCTCTCAAGCGCCTGATTTGTAAGAACTACTTCAAAAGAGGACCTGCTCGAAAT 1614
Db 2276 CACAAGGTAAGAGGACCTGCTGGCTGGAGCCCACTATACCAAGGAAGGCGCTGAGGGAAT 2335
Qy 1615 GATATTTACAGAACCAATGTACCGCACATCAGGTTTCAATTTAGAGAGATGATCTGGAGA 1674
Db 2336 GACATCCGCGGACCAATGTGCCAGATCCCGTGGGCTACCGCTACGAGACCTGTGC 2395
Qy 1675 AATGAATGAACCT 1688
Db 2396 CAGGAGCTGGCGCT 2409

RESULT 3
US-09-960-706-706
; Sequence 706, Application US/09960706
; Publication No. US20030134280A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William E.
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplasia
; FILE REFERENCE: 44921-5029-01US
; CURRENT APPLICATION NUMBER: US/09/960,706
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 60/223,323
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 09/873,319
; PRIOR FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 706
; LENGTH: 3386
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134280A1 M91029
US-09-960-706-706

Query Match 24.7%; Score 471.6; DB 13; Length 3386;
Best Local Similarity 58.3%; Pred. No. 2.9e-100;
Matches 847; Conservative 0; Mismatches 604; Indels 3; Gaps 1;

Qy 238 GGGAAACAAGAACTGCTGCCATAATCGGTTAAATCTTCTTGAGCATAAAGTTTAAATTC 297
Db 956 GGGCCCATAAAGTCAATCTGCTACCGCGGCTGCACTGAGCTCCCAAGTTCAGATG 1015
Qy 298 CATCTGATTTAAATGCGGATAGGAAATTTCTGCCGAGAGACTGCCCAACATCGTAT 357
Db 1016 CATGTGCTACTCAATGAGATGAAGGAGCTGCCGCCGAGAAAGTTCACACCGAGAT 1075
Qy 358 TTTTACAATGTCAGAGGTTGACATCATGTTTCAATTCATGATGATGATGAATCAAAAA 417
Db 1076 TTCTACAACATCCGAGGTGACACCAATCATGCTCTGCTGATGATGAACAGAG 1135
Qy 418 CATCTGTTGAGTTTCAAAAAATCAAAATCAAAAGAAAGAACCTGATGAGGTGGTCAATTTTC 477
Db 1136 CATCTGCTGCTTTCATCAAGCGGCAATGAAGCGCACCTGGAGGAGATCGTGACGCTG 1195
Qy 478 AGAGATGTTACTTATATGACCTTTAAAGAGGTTTTCAGAGCTTGGACTTAACCTGGGTAT 537
Db 1196 GAGCAGGCGCTGAAACAGACGCTCGGGAGGTCTTTGAGAGCATGAATCTCACGCGCTAC 1255
Qy 538 GATCTGAATGTTGATTTGCTAGATGTCATGACAGAAAGACATTTTCATCGTTTTCAC 597
Db 1256 GACCTGAGTGTGACACGCTGATGTCATCGGAGGAAACATTTTCCATCGCTTTGAC 1315
Qy 598 AAATTCATCTAAATACAAATCCATGTGGCCAAAGTAGGCTCAGAGAAATTTTCTCAAA 657

Db 1316 AAGTTTAAATGCCAAATACAAACCTATTGGGAGTCGGTCTCCGAGAGATCTTCATCAAG 1375
QY 658 CAGATAAATCTTATTCAAGGCGGTTTTCTTGCTGAGTTGACAAAGCAAGTTTTCTCTGAC 717
Db 1376 ACGGACAAACAGGAGTATCTGGGAAGTACTTTGCTACATCATCAAGGAGGTGATGTCAGAC 1435
QY 718 CTTTCTGCTAGCAAAATATCAGATGGCAGAAATATAGGATTTCAATCTACGGAAGGAAACAG 777
Db 1436 CTGGAGAGACAAATACAGAAATGAGAGCTGGGCTCTCCATTTACGGGCGCTCGAGG 1495
QY 778 AGTGAATGGACCAACTTGCAGTTGGATAGTGAACAATGAATTTGACAGTGGAAATGTT 837
Db 1496 GATGAGTGGACAAAGCTGGCGGCTGGCGGCTCATGCAACGGCTGCACTCCCCCAACGCTG 1555
QY 838 GTCTGCTGTTGAGATTTCAAGCTTATATATGTTGACAGGAATGGTATCGTTACA 897
Db 1556 CGTGGCTGGTGGAGTGGCGGCTCTTTGATGTACCGTACCAAGGGGCCAGCTGGGC 1615
QY 898 TCATTCCAAATCTCTTGACAAATTTTGGTCTCTTTTGGAGTTACTATTGATCA 957
Db 1616 AACTTCAGGAGATGCTGAGAACATCTTCTGCCACTGTTGAGGCCACTGTGCACCT 1675
QY 958 GCTTCACCCACAGCTCCATGCTTCTTGAAGCAGGTTGTAGGTTGGACCTGTTGAT 1017
Db 1676 GCCAGCCACCGCAACTGCATCTCTTCTTAGACAGCTGGATGTTTGGACAGCGTGGAT 1735
QY 1018 GATGAATGAACCAAGAGGGGTCACAAAGC---ACATGCCACACCTGACAGTGG 1074
Db 1736 GATGAGTCCAAAGCTGAAACCAATGCTTTCAACCTGGAGAGCCCTGCTGAGGCGTGG 1795
QY 1075 ACCAATGTGTTCAACCTGCTATTTTCATATTATGCTACTACTGCTATGCTAACTATTTC 1134
Db 1796 GTGAGGAGACAAACCCACTATGCTACTACTGCTATGCTACCTTTGGCCACATGGCC 1855
QY 1135 ACCCTAAACAGCTGGTGGAGTCAAAAGGGAATGACCACTATCAAAATCCCTCCACATGCT 1194
Db 1856 ATGTTGAACACCTGCGCAGCAGAGGGGCTTCCACACGTTTGTGTGAGGCCACACTGT 1915
QY 1195 GAGAGGCTGGAGATGTTGATGCTACTTGGCAGCAGCATTTCTCTGTCACACATATCA 1254
Db 1916 GGGGAGGCTGGGCCCCATCCACCACTGGTGTGACGCTTCAATGCTGCTGAGAACATTTCC 1975
QY 1255 CATGGAATTAATCTAAGGAAGTCTCTGTGCTTTCAGTACTTGTACTATCTGTGCTAGATT 1314
Db 1976 CAGGGCTCTTCTGCGCAAGGCCCTGCTGCTGAGTACTGCTGCTGCTGCTGCTGCTGCT 2035
QY 1315 GGTCTGGCGATGTCCTGCTGAGCAACACTCCTTATTTCTTGACTATCATCGCAACCTT 1374
Db 2036 GGCATCGCATGCTCCGCTCAGCAACACAGCCTCTTCTCAGCTATCACCGGAATCCG 2095
QY 1375 TTTCCAAAGTCTTCCAAAGAGTCTGAATGCTCTATATCTACGAGTACCTTTGGCAA 1434
Db 2096 CTACCGGAGTACCTGTCTCCGCGGCTCATGGTCTCCTGCTGCTGCTGCTGCTGCTGCTG 2155
QY 1435 ATTCACCTGACAAAGAACCAATGTTGGAAGAATACAGCAATTCGCTTCTGCTGCTGGAAG 1494
Db 2156 TTTCCATTTACCAAGGAGCGCTGATGGAGAGTACAGCATCCGACCCAGGCTGGAAG 2215
QY 1495 CTCAGTTCTTTGATTTATGCGAAATTTGCGAGAACTCTGTTTACCAATCTGGGTTTTCA 1554
Db 2216 CTCAGCTCTCGCATATGTTGAGCTGGCCCGCAACAGCGTGTCTATGAGCGGCTTCTCG 2275
QY 1555 CATGCTCTAAGGGGCTGAGTTGTAAGAACTATTTCAAAAGAGAACTGCTGCTGGAAT 1614
Db 2276 CACAAGGTAAAGAGCCACTGGCTGGGACCCAACTATACCAAGGAAGGCCCTGAGGGGAAT 2335
QY 1615 GATATTACAGAAACCAATGTACCGCACTCAGGGTTTCAATTTTATAGAGATGATCTGAGA 1674
Db 2336 GACATCGCGGACCAATGTGCCAGACATCCGGTGGGCTACCGCTACGAGACCTGTGTC 2395
QY 1675 AATGAAATGAAACT 1688

Db 2396 CAGGAGCTGGCGCT 2409
RESULT 4
US-10-119-926-83
; Sequence 83, Application US/10119926
; Publication No. US20030104413A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhao, Qing A.
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Xu, Chongjun
; APPLICANT: Yang, Yonghong
; APPLICANT: Xue, Aidong J.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104413A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 789CIP2B00N
; CURRENT APPLICATION NUMBER: US/10/119,926
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 83
; LENGTH: 5442
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2251)..(2820)
US-10-119-926-83
Query Match 24.0%; Score 458.8; DB 15; Length 5442;
Best Local Similarity 57.1%; Pred. No. 3.8e-97;
Matches 857; Conservative 0; Mismatches 642; Indels 3; Gaps 1;
QY 193 ACCTTCTTCACTGCTTGCATTTATTTCTCCGAGTAAACGGCTGCAGGGAACACAGAACT 252
Db 3913 ACCTTCTTTCAGAGATGATGAATTTTCTGCTTTAATGTCTCAAGGACCTGTTAAGACC 3972
QY 253 GTCTGCCATAATCGGTTAAATCTTCTGAGCATAAGTTTAAATTCATCTGATGTTAAAT 312
Db 3973 TATACCCACCGCGGCTGAAGTTCTCTCCCAAGTTCAGGTCATCAGATGCTTAAC 4032
QY 313 GCGGATAGGGAATTTCTTCCCGAAGACTGCCCAATCGTGATTTTACAAATGTCAGG 372
Db 4033 GAGATGGACGAGTTAAAGAGAGCTGAAAAAACCCCAACCGAGATTTTATAACTGCAGG 4092
QY 373 AAGTTTGACACTGCTTTCATCATTTCCAGATGCTGAATCAAAAACATCTGTTGAGGTTTC 432
Db 4093 AAGTTGGACACCATATCCATGACGCGCTTGCATGAAACCAACCAATCTGCTGCGTTT 4152
QY 433 ATAAAACTCAAACTAAGAAAGAACTGATGAGGTGGTCAATTTTCAGAGATGCTACTAT 492
Db 4153 ATTAGAATCTTACCAATTTGATGCTGACAGAGTGGTCTATAGCAACCAAGAGAAT 4212
QY 493 ATGACTTTAAAAGAGGTTTTTGAGAGCTTGGACTTAACTGGGTATGATCTGAATGTTGAT 552
Db 4213 CTGACCTTAAAGGAACTTTTCTGCTAAATTTAAAAATGCACTCTTATGACCTGCTGTTGAT 4272
QY 553 TTGCTAGATGCTCCATGACAGAAAAGACATTTTCATCGTTTTTGACAAAATTCATCTAAA 612
Db 4273 TCTCTGGATGTTTCATGCTGGACGCCACGCTTCCAGCGGTTTGAATAAGTTCAATGACAAA 4332


```

Db      243  TAAATAAGCTGATGAACAAATTTCCAAAGAAATCTTTACAAATATCTTATTTGCCATTATTG 184
Qy      941  AGGTTACTATTGATCCAGCTTTCACCCACAGCTCCATGCTCTTCTCGAAGCAGGTTGTAG 1000
Db      183  AAGTTACGAATGATCCAAATCGCATCTCGAACTACACAAGTTTCTTTCAGTTCTGTCGATAG 124
Qy      1001  GTTTGGACCTGGTTTGATGATGAAGTAAGTAAACCGAAGAGGGGTCCA---ACAAAGCACATGC 1057
Db      123  GTTTTGACTCGTGGGATGATGAAGTAAGTAAACCCGAAATCCATTATTTCGATNAGGATGTT 64
Qy      1058  CCACACTGAACAGTGGACCAATGTGTTCAACCTCGATTTTCATATTAATGCGTACTA 1115
Db      63  CGACACCAGAAGATGGACAGACACTGAAATCGCCCTATGTCATATTATCAATATTA 6

RESULT 9
US-09-960-352-3507
; Sequence 3507, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Mengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathalagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION
; FILE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 3507
; LENGTH: 438
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (350)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: 16-BOVMS1-004-Q1-E1-D11
US-09-960-352-3507

Query Match      4.6%; Score 87.2; DB 10; Length 438;
Best Local Similarity 53.0%; Pred. No. 2.2e-10;
Matches 185; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

Qy      682  TTTCTTGCTGAGTTGACAAGCAAGTTTCTCTGACCTTTCTGCTAGCAATATCAGATG 741
Db      4      TATTTTGCCACTATCATCAAGGAGGTAGGTGCAGACTTGGTGGAGGCCAAGTACCAGCAT 63
Qy      742  GCAGAATATAGGATTTTCAATCTACGGAAGGAACAGAGTGAATGGGACCACCTTGCAAGT 801
Db      64  GCGAGCCCCGCTTGTCCTATATGCGCCAGTCTCGAGGAGTGAGGCAAACTCTCCACC 123
Qy      802  TGGATAGTGAACAATGAATTGCACAGTGGAAATGTTGCTCTGGCTGGTTTCAGATTCCACGC 861
Db      124  TGGTTCTGTCMAAACCGCTCTACTGCCCCAACATGCATGCATGATGCCAGTCCCCAGG 183
Qy      862  TTATATAATGTGTACAAGGAATGGGTATCGTTTACATCATTTCCAAAATCTCTTCGACAC 921
Db      184  ATCTATGATGTGTTCGATCTAAGAACTTCTTCCACACTTTGGAAAGATGTTGGAGAT 243
Qy      922  ATTTTCGTTCTCTTTTTCAGGTTTACTATTGTATCCAGTTCACACCCACAGCTCCCATGTC 981
Db      244  GTTTTCATGCGAGTGTTCGAGGCCACCATCAACCTTCAGGCTCACCCAGACCTCAGTGTT 303
Qy      982  TTCCTCAAGCAGGTTGTAGGTTGGACCTGGTTGATGATGAAGTAAAC 1030
Db      304  TTCCTCAAGCACATTACAGGCTTCGACAGTGTAGATGATGAGTCCANAC 352

RESULT 10
US-09-930-213-630/c
; Sequence 630, Application US/09930213

```

Publication No. US20030170625A1
GENERAL INFORMATION:
APPLICANT: ROSENTHAL, ANDRE
APPLICANT: HINZMANN, BERND
APPLICANT: SCHAFER, REINHARD
APPLICANT: ZUBER, JOHANNES
APPLICANT: TCHE-NITSE, OLEG
APPLICANT: GRIPS, MARTIN
APPLICANT: HELNEGEL, MARTIN
APPLICANT: SCHMITZ, ANNE-CHANTAL
APPLICANT: SEPS, CHRISTINE
TITLE OF INVENTION: DETECTION OF DIFFERENTIAL GENE EXPRESSIONS
FILE REFERENCE: ALBRE-14
CURRENT APPLICATION NUMBER: US/09/930,213
CURRENT FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: DE 10004102.7
PRIOR FILING DATE: 2000-01-31
NUMBER OF SEQ ID NOS: 885
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 630
LENGTH: 255
TYPE: DNA
ORGANISM: Rattus sp.
FEATURE: modified base
NAME/KEY: (7)
LOCATION: (7)
OTHER INFORMATION: a, t, c, g, other or unknown
US-09-930-213-630

Query Match 4.0%; Score 77.2; DB 13; Length 255;
Best Local Similarity 65.4%; Pred. No. 3.7e-08;
Matches 106; Conservative 4; Mismatches 52; Indels 0; Gaps 0;

Qy 1749 TCGTGTATACAGACGAGTGGCTGTAGCTGCTATGGGAATTACTTCTATGTTTGGT 1808
Db 198 TGTGTTTCCCTCCCTCTTTTATCGTTTGGGTTTATTTTCTTTTGTGCGC 139

Qy 1809 ATGCTTTCTTATCTATGCGAAATCAACTCGAACTTCAAAAAA 1868
Db 138 ATTATAMCCTTTTCTTTTAAATTTCCCTTTTGGGAAAAA 79

Qy 1869 AAAAAA 1910
Db 78 AAAAAA 37

RESULT 11
US-10-063-735-17
Sequence 17, Application US/10063735
Publication No. US2003013882A1
GENERAL INFORMATION:
APPLICANT: Eaton, Dan L.
APPLICANT: Filvaroff, Ellen
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3230R1C1
CURRENT APPLICATION NUMBER: US/10/063,735
CURRENT FILING DATE: 2002-05-08
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 170
SEQ ID NO 17
LENGTH: 1672
TYPE: DNA
ORGANISM: Homo Sapien
US-10-063-735-17

Query Match 4.0%; Score 77.2; DB 13; Length 1672;
Best Local Similarity 72.5%; Pred. No. 9.9e-08;
Matches 100; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

Qy 1773 TGTAGCTGCTATGGGAATTACTTCTATGTTTGGTATGCTTTCTTATCTATGCGAAAT 1832
Db 1525 TTTTGCTAATTTGTTACTTTTCTTGTCTAATTTGGAGATTAACTCATTTTAAATAAAA 1584

Qy 1833 TCAACTTCGAACTTCAAAAAA 1910
Db 1585 TTATGCTAGATTAAAAA 1662

Qy 1893 AAAAAA 1910
Db 1645 AAAAAA 1662

RESULT 13
US-10-063-586-17
Sequence 17, Application US/10063586
Publication No. US2003017668A1
GENERAL INFORMATION:
APPLICANT: Eaton, Dan L.
APPLICANT: Filvaroff, Ellen
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.

Query Match 4.0%; Score 77.2; DB 13; Length 1672;
Best Local Similarity 72.5%; Pred. No. 9.9e-08;
Matches 100; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

Qy 1773 TGTAGCTGCTATGGGAATTACTTCTATGTTTGGTATGCTTTCTTATCTATGCGAAAT 1832
Db 1525 TTTTGCTAATTTGTTACTTTTCTTGTCTAATTTGGAGATTAACTCATTTTAAATAAAA 1584

Qy 1833 TCAACTTCGAACTTCAAAAAA 1910
Db 1585 TTATGCTAGATTAAAAA 1662

Qy 1893 AAAAAA 1910
Db 1645 AAAAAA 1662

Publication No. US20030170625A1
GENERAL INFORMATION:
APPLICANT: ROSENTHAL, ANDRE
APPLICANT: HINZMANN, BERND
APPLICANT: SCHAFER, REINHARD
APPLICANT: ZUBER, JOHANNES
APPLICANT: TCHE-NITSE, OLEG
APPLICANT: GRIPS, MARTIN
APPLICANT: HELNEGEL, MARTIN
APPLICANT: SCHMITZ, ANNE-CHANTAL
APPLICANT: SEPS, CHRISTINE
TITLE OF INVENTION: DETECTION OF DIFFERENTIAL GENE EXPRESSIONS
FILE REFERENCE: ALBRE-14
CURRENT APPLICATION NUMBER: US/09/930,213
CURRENT FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: DE 10004102.7
PRIOR FILING DATE: 2000-01-31
NUMBER OF SEQ ID NOS: 885
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 630
LENGTH: 255
TYPE: DNA
ORGANISM: Rattus sp.
FEATURE: modified base
NAME/KEY: (7)
LOCATION: (7)
OTHER INFORMATION: a, t, c, g, other or unknown
US-09-930-213-630

Query Match 4.0%; Score 77.2; DB 13; Length 255;
Best Local Similarity 65.4%; Pred. No. 3.7e-08;
Matches 106; Conservative 4; Mismatches 52; Indels 0; Gaps 0;

Qy 1749 TCGTGTATACAGACGAGTGGCTGTAGCTGCTATGGGAATTACTTCTATGTTTGGT 1808
Db 198 TGTGTTTCCCTCCCTCTTTTATCGTTTGGGTTTATTTTCTTTTGTGCGC 139

Qy 1809 ATGCTTTCTTATCTATGCGAAATCAACTCGAACTTCAAAAAA 1868
Db 138 ATTATAMCCTTTTCTTTTAAATTTCCCTTTTGGGAAAAA 79

Qy 1869 AAAAAA 1910
Db 78 AAAAAA 37

RESULT 11
US-10-063-735-17
Sequence 17, Application US/10063735
Publication No. US2003013882A1
GENERAL INFORMATION:
APPLICANT: Eaton, Dan L.
APPLICANT: Filvaroff, Ellen
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3230R1C1
CURRENT APPLICATION NUMBER: US/10/063,735
CURRENT FILING DATE: 2002-05-08
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 170
SEQ ID NO 17
LENGTH: 1672
TYPE: DNA
ORGANISM: Homo Sapien
US-10-063-735-17

Query Match 4.0%; Score 77.2; DB 13; Length 1672;
Best Local Similarity 72.5%; Pred. No. 9.9e-08;
Matches 100; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

Qy 1773 TGTAGCTGCTATGGGAATTACTTCTATGTTTGGTATGCTTTCTTATCTATGCGAAAT 1832
Db 1525 TTTTGCTAATTTGTTACTTTTCTTGTCTAATTTGGAGATTAACTCATTTTAAATAAAA 1584

Qy 1833 TCAACTTCGAACTTCAAAAAA 1910
Db 1585 TTATGCTAGATTAAAAA 1662

Qy 1893 AAAAAA 1910
Db 1645 AAAAAA 1662

RESULT 13
US-10-063-586-17
Sequence 17, Application US/10063586
Publication No. US2003017668A1
GENERAL INFORMATION:
APPLICANT: Eaton, Dan L.
APPLICANT: Filvaroff, Ellen
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.

Query Match 4.0%; Score 77.2; DB 13; Length 1672;
Best Local Similarity 72.5%; Pred. No. 9.9e-08;
Matches 100; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

Qy 1773 TGTAGCTGCTATGGGAATTACTTCTATGTTTGGTATGCTTTCTTATCTATGCGAAAT 1832
Db 1525 TTTTGCTAATTTGTTACTTTTCTTGTCTAATTTGGAGATTAACTCATTTTAAATAAAA 1584

Qy 1833 TCAACTTCGAACTTCAAAAAA 1910
Db 1585 TTATGCTAGATTAAAAA 1662

Qy 1893 AAAAAA 1910
Db 1645 AAAAAA 1662


```
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,586
; CURRENT FILING DATE: 2002-05-03
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 17
; LENGTH: 1672
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-063-586-17

Query Match      4.0%; Score 77.2; DB 13; Length 1672;
Best Local Similarity 72.5%; Pred. No. 9.9e-08;
Matches 100; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 1773 TGTAGCTGCTATGGGAATTATACCTTCATGTTTGGTATGCTTTCCTTATCTATGCGCAAT 1832
Db 1525 TTTTGCTAATTGTTACTTTTCTTTGCTAATTGGAAGATTAACTCACTTTTAAATAAAA 1584

QY 1833 TCAACTTCGAACCTTCAAAAAA 1910
Db 1585 TTATGCTCTAAGATTAAAAA 1662

QY 1893 AAAAAA 1910
Db 1645 AAAAAA 1662

RESULT 14
US-10-063-510-17
; Sequence 17, Application US/10063510
; Publication No. US20030180837A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,510
; CURRENT FILING DATE: 2002-05-01
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 17
; LENGTH: 1672
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-063-510-17

Query Match      4.0%; Score 77.2; DB 13; Length 1672;
Best Local Similarity 72.5%; Pred. No. 9.9e-08;
Matches 100; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 1773 TGTAGCTGCTATGGGAATTATACCTTCATGTTTGGTATGCTTTCCTTATCTATGCGCAAT 1832
Db 1525 TTTTGCTAATTGTTACTTTTCTTTGCTAATTGGAAGATTAACTCACTTTTAAATAAAA 1584

QY 1833 TCAACTTCGAACCTTCAAAAAA 1910
Db 1585 TTATGCTCTAAGATTAAAAA 1662

QY 1893 AAAAAA 1910
Db 1645 AAAAAA 1662

RESULT 15
US-10-063-514-17
; Sequence 17, Application US/10063514
; Publication No. US20030181707A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,514
; CURRENT FILING DATE: 2002-05-01
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 17
; LENGTH: 1672
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-063-514-17

Query Match      4.0%; Score 77.2; DB 13; Length 1672;
Best Local Similarity 72.5%; Pred. No. 9.9e-08;
Matches 100; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 1773 TGTAGCTGCTATGGGAATTATACCTTCATGTTTGGTATGCTTTCCTTATCTATGCGCAAT 1832
Db 1525 TTTTGCTAATTGTTACTTTTCTTTGCTAATTGGAAGATTAACTCACTTTTAAATAAAA 1584

QY 1833 TCAACTTCGAACCTTCAAAAAA 1910
Db 1585 TTATGCTCTAAGATTAAAAA 1662

QY 1893 AAAAAA 1910
Db 1645 AAAAAA 1662

Search completed: December 15, 2003, 22:56:48
Job time : 609 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003, Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2003, 16:43:05 ; Search time 3666 Seconds
(without alignments)
12662.725 Million cell updates/sec

Title: US-10-019-633-1

Perfect score: 1910

Sequence: 1 ccacgcctcgggtgtctcc.....aaaaaaaaaaaaaaaaaaaaa 1910

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1910	100.0	1911	11	AY109001 Zea mays
2	1067.8	55.9	2196	11	AY107533 Zea mays
3	682.8	35.7	832	14	CB642483 OSJNEB02K
4	651.4	34.1	798	14	CB643461 OSJNEB04C

5	606	31.7	624	9	AW017903	AN017903 614020B05
6	590.8	30.9	877	14	CB660024	CB660024 OSJNEB01H
7	576.8	30.2	634	9	AW671750	AW671750 LG1 351.F
8	555.4	29.1	845	14	CB668055	CB668055 OSJNEB15L
9	551.2	28.9	757	14	CB671812	CB671812 OSJNEB05I
10	517	27.1	872	12	BM359276	BM359276 GA_Ea001
11	508.2	26.6	647	13	BU973129	BU973129 HB23JN17x
12	507.6	26.6	676	12	BI955474	BI955474 HVSME002
13	494.6	25.9	880	10	BF268586	BF268586 GA_EB000
14	490.6	25.7	913	10	BG365787	BG365787 HVSME1000
15	488.2	25.6	909	10	BG442858	BG442858 GA_Ea001
16	487.8	25.5	3586	11	AK004759	AK004759 Mus_muscu
17	460	24.1	676	13	BQ871213	BQ871213 QG11D06.
18	456	23.9	954	10	BF268791	BF268791 GA_EB000
19	453.2	23.7	546	13	BU991366	BU991366 HD06L13r
20	445.2	23.3	658	14	CD226569	CD226569 CCL1 46 B
21	444	23.2	652	14	CA917577	CA917577 EST641724
22	443.4	23.2	461	12	BM501264	BM501264 PAC000000
23	442.2	23.2	679	13	BQ96927	BQ96927 QG14D05.
24	435.6	22.8	519	12	BI948901	BI948901 HVSME1001
25	434	22.7	622	10	AW906456	AW906456 EST342578
26	429.6	22.5	524	10	BE425619	BE425619 WHE0322 C
27	429	22.5	3402	11	AK034844	AK034844 Mus_muscu
28	415	21.7	666	12	BG86862	BG86862 EST512713
29	412	21.6	955	12	EM084714	EM084714 PPAP_10_E
30	411.2	21.5	759	10	BE034209	BE034209 MH01C02_M
31	410.8	21.5	647	13	BQ465687	BQ465687 HU04G08r
32	410.2	21.5	652	14	CB869343	CB869343 HC11A23w
33	409	21.4	803	14	CA767547	CA767547 AF53-Rpf
34	406.6	21.3	893	10	BF268580	BF268580 GA_EB000
35	404.2	21.2	580	10	BE475206	BE475206 sp73e06.Y
36	403.4	21.1	566	12	BI785902	BI785902 sai30e01.
37	402	21.0	426	9	AW287505	AW287505 LG1 228 H
38	401.6	21.0	645	13	BQ994451	BQ994451 QGF7E08.Y
39	400	20.9	860	12	BM359287	BM359287 GA_Ea001
40	396.2	20.7	583	12	BG890464	BG890464 EST516315
41	396	20.7	588	10	BG648222	BG648222 EST509841
42	392.8	20.6	646	14	CA840318	CA840318 MCT036B07
43	383.4	20.1	589	10	BE362865	BE362865 DGI 90 C0
44	377.6	19.8	577	13	BQ778755	BQ778755 946114A12
45	371.4	19.4	533	9	AW831078	AW831078 sm08a08.Y

ALIGNMENTS

RESULT 1	AY109001	AY109001	1911 bp	mRNA	linear	HTC 17-OCT-2002
LOCUS	Zea mays PC0134657 mRNA sequence.					
DEFINITION	Zea mays PC0134657 mRNA sequence.					
ACCESSION	AY109001					
VERSION	AY109001.1 GI:21212355					
KEYWORDS	HTC.					
SOURCE	Zea mays					
ORGANISM	Zea mays					
REFERENCE	Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.					
AUTHORS	Hayney, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whittsitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.					
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes					
JOURNAL	Unpublished (2002)					
REFERENCE	2 (bases 1 to 1911)					
AUTHORS	Coe, E.H.					
TITLE	Direct Submission					
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA					
COMMENT	If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemdb.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the					

maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB:

FEATURES
source

Location/Qualifiers
1..1911
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:637880"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Cornsensus Library"

/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

BASE COUNT 597 a 391 c 383 g 540 t

ORIGIN

Query Match 100.0%; Score 1910; DB 11; Length 1911;
Best Local Similarity 100.0%; Pred. No. 6.9e-176;
Matches 1910; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCACGGCTCCGGTTCCTCCATGGGAGAGGAGGTCATAAATGACCCCTGTAATCCAAA 60
DB 1 CCACGGCTCCGGTTCCTCCATGGGAGAGGAGGTCATAAATGACCCCTGTAATCCAAA 60

QY 61 CTTAAACCCCAACCCGTTCACTTATGTGCTGAAACCAAGTCAGAGCATGTTTCCAAACT 120
DB 61 CTTAAACCCCAACCCGTTCACTTATGTGCTGAAACCAAGTCAGAGCATGTTTCCAAACT 120

QY 121 GTTGATGGCTTATCCATGTTTATGCGGATAAAGATTTGACGAGAGCATTTATCTGTG 180
DB 121 GTTGATGGCTTATCCATGTTTATGCGGATAAAGATTTGACGAGAGCATTTATCTGTG 180

QY 181 GCTGATGCTACAACTTCTTCACTGACTTGCATTATTTCTCCGAGTAACGGCTGCAGGG 240
DB 181 GCTGATGCTACAACTTCTTCACTGACTTGCATTATTTCTCCGAGTAACGGCTGCAGGG 240

QY 241 AACACAAGAACTGTGCAATATCGTTTAAATCTTTTGAGCAATAAGTTTAAATTCAT 300
DB 241 AACACAAGAACTGTGCAATATCGTTTAAATCTTTTGAGCAATAAGTTTAAATTCAT 300

QY 301 CTGATGTTAAATGCGGATAGGAAATTTCTGCGCAGAGACATGCCCAATCGTATTT 360
DB 301 CTGATGTTAAATGCGGATAGGAAATTTCTGCGCAGAGACATGCCCAATCGTATTT 360

QY 361 TACAATGTCAGGAGGTTGACACTCATGTTTCATCTTACGATGATGAATCAAAACAT 420
DB 361 TACAATGTCAGGAGGTTGACACTCATGTTTCATCTTACGATGATGAATCAAAACAT 420

QY 421 CTGTTGAGGTTTCAATAAATCCAAACTTAAGAAAGAACCTGATGAGTGGTTCATTTT 480
DB 421 CTGTTGAGGTTTCAATAAATCCAAACTTAAGAAAGAACCTGATGAGTGGTTCATTTT 480

QY 481 GATGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 481 GATGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540

QY 541 CTGAATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGAT 600
DB 541 CTGAATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGATGAT 600

QY 601 TTCAATCTTAAATFACAAATCCAACTTAAGAAAGAACCTGATGAGTGGTTCATTTT 660
DB 601 TTCAATCTTAAATFACAAATCCAACTTAAGAAAGAACCTGATGAGTGGTTCATTTT 660

QY 661 GATAATCTTAAATCAAGGCGGTTTCTGCTGATGATGATGATGATGATGATGATGATGAT 720
DB 661 GATAATCTTAAATCAAGGCGGTTTCTGCTGATGATGATGATGATGATGATGATGATGAT 720

QY 721 TCTGCTAGCAATATACAGATGGCAGAAATATAGGATTTCAATCTACGGAAGGAAACAGAT 780

DB 721 TCTGCTAGCAATATACAGATGGCAGAAATATAGGATTTCAATCTACGGAAGGAAACAGAT 780

QY 781 GAATGGGACCAACTTGAAGTTGATAGTGAACAAATGAATTTGACAGTGGAAATGTTGTC 840
DB 781 GAATGGGACCAACTTGAAGTTGATAGTGAACAAATGAATTTGACAGTGGAAATGTTGTC 840

QY 841 TGGCTGGTTTCAGATTCACAGCTTATATAATGTGTACAAAGAAATGGGTATCGTTACATCA 900
DB 841 TGGCTGGTTTCAGATTCACAGCTTATATAATGTGTGTACAAAGAAATGGGTATCGTTACATCA 900

QY 901 TTCCAAATCTTCTGTGACAACTTTCTCTCTCTTTTGAGTTTACATTTGATCCAGCT 960
DB 901 TTCCAAATCTTCTGTGACAACTTTCTCTCTCTTTTGAGTTTACATTTGATCCAGCT 960

QY 961 TCACACCAAGCTCCATGTTCTCTGACAGAGTGTGTAGGTTGGACCTGGTTGATGAT 1020
DB 961 TCACACCAAGCTCCATGTTCTCTGACAGAGTGTGTAGGTTGGACCTGGTTGATGAT 1020

QY 1021 GAAAGTAAACCAAGAGGCTCCCAACCAAGCACATGCCACACCTGAACAGTGGACCAAT 1080
DB 1021 GAAAGTAAACCAAGAGGCTCCCAACCAAGCACATGCCACACCTGAACAGTGGACCAAT 1080

QY 1081 GTGTTCAACCTGCAATTTTCAATTTATGCGTACTGCTATGCTTAATTTCAACCTTA 1140
DB 1081 GTGTTCAACCTGCAATTTTCAATTTATGCGTACTGCTATGCTTAATTTCAACCTTA 1140

QY 1141 AACAGCTGCTGAGTCAAGGGAATGACCACTATCAAAATTCGTCACATGCTGGAGAG 1200
DB 1141 AACAGCTGCTGAGTCAAGGGAATGACCACTATCAAAATTCGTCACATGCTGGAGAG 1200

QY 1201 GCTGAGATGTTGATCACTTGGCAGCGACATTTCTTCTGTCTCAACATATCACATGGA 1260
DB 1201 GCTGAGATGTTGATCACTTGGCAGCGACATTTCTTCTGTCTCAACATATCACATGGA 1260

QY 1261 ATTAATCTAAAGAAAGTCTCTCTGCTTCAAGTACTTGTACTATCTTGGTCAATTTGCTG 1320
DB 1261 ATTAATCTAAAGAAAGTCTCTCTGCTTCAAGTACTTGTACTATCTTGGTCAATTTGCTG 1320

QY 1321 GCGATGCTCCCATTTGAGCAACAACTCTTATTTCTGACTATCATCGCAACCTTTTCCA 1380
DB 1321 GCGATGCTCCCATTTGAGCAACAACTCTTATTTCTGACTATCATCGCAACCTTTTCCA 1380

QY 1381 ACCTTCTTCAACAGAGTCTGAATGCTCTATCTACGAGTACCTTTGCAAAATTCAC 1440
DB 1381 ACCTTCTTCAACAGAGTCTGAATGCTCTATCTACGAGTACCTTTGCAAAATTCAC 1440

QY 1441 CTGCAAAAGAAACCAATTTGGTGGAAAGATPACAGCAATGCTGCTTCTGCTGGAAGCTCAGT 1500
DB 1441 CTGCAAAAGAAACCAATTTGGTGGAAAGATPACAGCAATGCTGCTTCTGCTGGAAGCTCAGT 1500

QY 1501 TCTTGTGATTTATGGGAAATTTGCGAGGAACTCTGTTTACCAATCTGGGTTTTCATGCT 1560
DB 1501 TCTTGTGATTTATGGGAAATTTGCGAGGAACTCTGTTTACCAATCTGGGTTTTCATGCT 1560

QY 1561 CTCAGGCGCACTGGATTTGGAAGAACTACTTTCAAAAGAGGACCTGCTGGAAATGATATT 1620
DB 1561 CTCAGGCGCACTGGATTTGGAAGAACTACTTTCAAAAGAGGACCTGCTGGAAATGATATT 1620

QY 1621 CACGAAACCAATGTACCCGACATCAGGGTTCAATTTTAGAGAGATGATCTGGAGAAATGAA 1680
DB 1621 CACGAAACCAATGTACCCGACATCAGGGTTCAATTTTAGAGAGATGATCTGGAGAAATGAA 1680

QY 1681 ATGAACTAGTGTACTCTGACAAATGAGATCTTAATACCAGACGAGCTGGACCTGTGAAGAT 1740
DB 1681 ATGAACTAGTGTACTCTGACAAATGAGATCTTAATACCAGACGAGCTGGACCTGTGAAGAT 1740

QY 1741 GTCCAGCCTCGGTATACAGACGAGTTCGTTGCTGTAGCTGTATGGGAATTTATCTTCAT 1800
DB 1741 GTCCAGCCTCGGTATACAGACGAGTTCGTTGCTGTAGCTGTATGGGAATTTATCTTCAT 1800

QY 1801 GTTTGGTATGCTTTCTTATCTATGGCAAAATTTCAACTTCGAACCTTCAAAAAA 1860

```

Db      1801 GTTTGGTATGCTTTCCTTATCTATGGCAATCAACTTCGAACTTCAAAAA 1860
QY      1861 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1910
Db      1861 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1910

RESULT 2
LOCUS   AY107533
DEFINITION Zea mays PC0130446 mRNA sequence.
ACCESSION AY107533
VERSION   AY107533.1 GI:21210611
KEYWORDS HTC.
SOURCE   Zea mays
ORGANISM Zea mays
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
          clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 2196)
          Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitesitt,M.S.,
          Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
          Maize Mapping Project/DuPont Consensus Sequences for Design of
          Overgo Probes
          Unpublished (2002)
REFERENCE 2 (bases 1 to 2196)
          Coe,E.H.
          Direct Submission
          Submitted (25-APR-2002) Maize Mapping Project, University of
          Missouri, Columbia, MO 65211, USA
          If you are interested in getting corresponding physical clones,
          these are publicly available from ZmDB and may be found by BLAST
          searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
          www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
          maize cDNA sequences is either Virginia Walbot, Stanford or Pat
          Schnable, Iowa State, then clones may be requested from ZmDB:
          www.zmdb.iastate.edu.

FEATURES             Location/Qualifiers
     source           1..2196
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /db_xref="MaizeDB:637602"
                     /db_xref="taxon:4577"
                     /clone_lib="Maize Mapping Project/DuPont Consensus
                     Library"
     note             "this sequence is part of a project of EST
                     assemblies resulting from the application of public
                     contigs to seed Dupont contigs; this resource was
                     assembled by Dupont as part of a collaboration for the
                     overgo addressing of BACs in conjunction with the Maize
                     Mapping Project"
BASE COUNT          655 a 442 c 478 g 621 t
ORIGIN

Query Match          55.9%; Score 1067.8; DB 11; Length 2196;
Best Local Similarity 76.5%; Pred. No. 9.7e-95;
Matches 1309; Conservative 0; Mismatches 402; Indels 0; Gaps 0;

QY      13 GTTGCTCCATGGAGAGGAGGTGATAAATGACCCCTGTACTCCAAAACCTAACCCCAAC 72
Db      134 GTTGCTCCCTGGGAAAAAGAGATCAATACTGATCTCTAGTACTCCAAAGCCCTAATCCNAAT 193

QY      73 CGGTTCACTATGTGCTGCAACCAAGTCAGAGCATGTTTTCCAACTGTTGATGGCGTT 132
Db      194 CGGTTCACTATGAACATCAGCTAAACCGATCAACATTTGCAATGTTGATGGCGTC 253

QY      133 ATCCATGTTTATGCGGATAAGATTGTACGGAGAGCAATTAATCCTGTGGCTGATGCTACA 192
Db      254 ATTCATGTGTACCCCAATAAAGATTCTAAAGAAAGACTCTATCCAGTTGCCGATCAACT 313

QY      193 ACCTTCTTCACGACTTGCAATATATTTCTCCGAGTAAAGGCTGACAGGAACACAAGAAGT 252
Db      314 ACCTTCTTTACTGATGATGCACTACGCTCTTCTGTGTGGTGGCTGGGGGATATCCGAACC 373

```

```

QY      253 GTCTGCCATAATCGGTTAAATCTTTCTGAGCATAAAGTTTAAATTCATCTCTGATGTTAAAT 312
Db      374 GTGTGTCATCACCGTTTAAATCTTTCTAGAGCAAAAATTTAAATCTTCACTTGTGTCAT 433

QY      313 CGGATAGGGAAATTTCTTGCCAGAGACTGCCCCACATCGTGATTTTAAATGTCAGG 372
Db      434 ACGGATAGAGAACTGCTTGTCTCAGAAAGCTGCACCTCACCGGAGCTTCTCAATGTTAGG 493

QY      373 AAGGTTGACACTCATGTTTCATCTCAGCATGCAATCAAAAACATCTCTGTGAGGTTTC 432
Db      494 AAGGTTGATACTCATGTTTCATCTCAGCATGCAATGACCAAGAGCACCCTGTTGAGATT 553

QY      433 ATAAATCCAACTAAAGAAAAAGAAACCTGATCAGGTGGTCATTTTTCAGAGATGGTACTTAT 492
Db      554 ATCAAGCTTAAGTTGAGGAAGAAACCTGATCAGGTGTTAACTCTTTAGAGATGGGACCTAT 613

QY      493 ATGACTTTAAAGAGGTTTTTTCAGAGCTTTCAGCTTAACCTGGGTATGATCTGAATGTTGAT 552
Db      614 TTGACTCTTAAGGAGGTTTTTTCAGAGTTTGGATTTGACAGGGTATGACCTCAATGTTGAT 673

QY      553 TTGCTAGATGCCATGACAGACAAAAGACATTTTCATCGTTTTTGACAAATTCATCTCAAAA 612
Db      674 CTCTTAGATGTTTCATGCTGATAGATACATTTTCATCGTTTTTGACAAATTCATCTCAAAA 733

QY      613 TACAATCCATGTGGCCAAAGTAGGCTCAGAGAAATTTTCTCTCAACAGAGATAATCTTAT 672
Db      734 TATAATCCATGTGTCATCTCGCTGAGGAGATATTTTAAAGCAGGATAATCTTAT 793

QY      673 CAAGCCGTTTTCTGCTGAGTTGACAAAGCAAGTTTTCTCTGACCTTTCTGCTAGCAAAA 732
Db      794 CAAGGTCGCTTCTCTGCTGAGCTGACAAAAGAGATTTTCTGACCTTTGAAAGCAAGTAAA 853

QY      733 TATCAGATGGCAGAAATATAGGATTTCAATCTACCGAAGGAAACAGAGTGAATGGAGCCAA 792
Db      854 TATCAGATGGCTGAGTATCGGATATCTATTTATGGAGGAAAGAAAGCAATGGAGCCAA 913

QY      793 CTTCAGAGTTGGATVAGTGAAACAATGTCACAGTGGAAATGTTGTCTGGCTGGTTCCAG 852
Db      914 ATGGCAAGTTGGATGAGTGAACAATGAATTTATAGCAGAGATGTTGTTGGTTAAATTCAG 973

QY      853 ATTCACGCTTATATATATGTTGTAAGAAATGGGTATCGTTATCATCATCTTCCAAATCTT 912
Db      974 ATTCCTCGGATATACAATGTATACAGGAGATGGGCACAATCAATTTCTTCCAAAACCTC 1033

QY      913 CTTCACAACTTTTTCGTTCTCTCTTTTGGAGTTACTATGATCCAGCTTTCACACCCACAG 972
Db      1034 CTTCACAAATTTTTCCTCTCTTTTATGAAGTAACTGTTGACCCAGCTTTCACATCCTCAG 1093

QY      973 CTCCATGTTCTTCTGAAAGCAGGTTGTAGGGTTGGAGCTGGTTGATGATGAAGTAAACCA 1032
Db      1094 CTCCATGTTTCTTAAACACAGTTGTTGGTTGGATTTGGTGGATGATGAAGCAACCT 1153

QY      1033 GAAAGGGCTCAACAAAGCAGCATGCCCAACCTGAAACAGTGGACCAATGTGTTCAACCT 1092
Db      1154 GAAAGAGCCCAACAAACCAACATGCCAACACCTGAGCAATGAGCAATAATTTTCAATCCA 1213

QY      1093 GCATTTTCATATTAATGATGTTACTGCTATGCTAACTTATTCACCTCAACCAAGCTGGT 1152
Db      1214 GCCTATGGTATTAATGTTACTATTTGCTATGCCAATCTGTATACACTGAACCAAGCTTGGT 1273

QY      1153 GAGTCAAGGGGAATGACCACTATCAAAATTCGTCACATGCTGGAGAGGCTGGAGATGTT 1212
Db      1274 GAGTCCAAGGCGATGACAAACAATCAAACTTCGTCCAACCTGCGGGAGGCTGGTGATATC 1333

QY      1213 GATCACTGGCAGCGCAATTTCTTCTCTGTGCACAAATATACATGGAATTAATCTAAGG 1272
Db      1334 GATCACCTTGTGTCAGCATTTCTTACTTCAATAACATTTGCTCATGGGTGAACCTAAAG 1393

QY      1273 AAGTCTCTGTGCTTTCAGTACTTGTACTATCTTGGTCAGATTGCTGTGGCGATGTCGCCA 1332
Db      1394 AATCCCTCTGACTTCACTGACCTGATTTATTTATCTTGTCTGAGATTGTTGCTAGCAATGTCCTCC 1453

```

QY 1333 TTGAGCAACAACCTCTTATTTCTTGACTATCATCGCAACCCCTTTTCCAAAGTCTTCTCAA 1392
 Db 1454 TTAAGTAACAACCTCACTGTTTATCGACTATCATCGGAACCCCTTCCCAACATTTTCTTA 1513
 QY 1393 CGAGGTCTGAATGTCTCATATCTAGGATGAGCCCTTTGCAAAATTCACCTGACAAAGAA 1452
 Db 1514 AGGGGTCTTAAGGTGCTCTTTGCACTGACGACCCCTCTACAAATTCACCTGACGAGGAA 1573
 QY 1453 CCAATGGTGGAGAAATACAGCATTTGCTGCTGCGTGGAGCTCAAGTCTTGTGATTTA 1512
 Db 1574 CTTTGGTGGAGAAATACAGTGTGCTGCTTCACTTGGAGCTGAGCTCATGTGACTG 1633
 QY 1513 TGGAAATTCGAGGAACTCTGTTTACCAATCTGGGTTTTCATCATGCTCTCAAGGCGCAC 1572
 Db 1634 TGTGAATTCGTAGAAATCTGTCTACCAATCGGATTCCTCACAGGCTCAAGTCTCAC 1693
 QY 1573 TGGATTTGTAAGAACTACTTCAAAGAGGACCTGCTGGAAATGATATTCACAGAACCAAT 1632
 Db 1694 TGGATCGGGAGAACTACTACAGAGGGGTCCGGAGGCAATGACATCCACCAGACAAAT 1753
 QY 1633 GTACCCGACATCAGGCTTCAATTTAGAGAGATGATCTGGAGAAATGAAATGAACTAGTG 1692
 Db 1754 GTTCTTCATCATCAGGTTGAATTTTCACACATCTATCTGGAAGAGAAATGAGCTGATA 1813
 QY 1693 TACTCTGACAATGAGATCTTAATACCAGACG 1723
 Db 1814 CATTTTGGAAATGTTAACTACCAGGAAG 1844

RESULT 3
 CB642483
 LOCUS
 DEFINITION OSJNEB02K24. f OSJNEB Oryza sativa (japonica cultivar-group) cDNA
 clone OSJNEB02K24 5', mRNA sequence.
 CB642483
 CB642483.1 GI:29637474
 EST.

ORGANISM
 SOURCE Oryza sativa (japonica cultivar-group)
 Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,B.,
 Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang G.
 Large-scale identification of ESTs involved in the interaction
 between rice and Magnaporthe grisea
 Unpublished
 Contact: Rod Wing
 Arizona Genomics Institute
 University of Arizona
 Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
 85721-0088, USA
 Tel: 520 626 3967
 Fax: 520 621 9288
 Email: http://genome.arizona.edu

PCR Primers
 FORWARD: gta aaa cga cgg cca gtc
 BACKWARD: gga aac agc tat gac cat g
 Plate: 02 row: K column: 24
 Seq primer: gta aaa cga cgg cca gtc.
 Location/Qualifiers
 1..832
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="OSJNEB02K24"
 /tissue_type="Leaf"
 /dev_stage="3 week"
 /lab_host="DH10B"
 /clone_lib="OSJNEB"
 /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:

FEATURES
 source
 1..832
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="OSJNEB02K24"
 /tissue_type="Leaf"
 /dev_stage="3 week"
 /lab_host="DH10B"
 /clone_lib="OSJNEB"
 /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:

BASE COUNT 245 a 169 c 169 g 249 t
 ORIGIN
 Query Match 35.7%; Score 682.8; DB 14; Length 832;
 Best Local Similarity 88.9%; Pred. No. 2.1e-57;
 Matches 738; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
 QY 157 TGTACGGAGAGCATTTATCTCTGCTGATGCTACAACTCTTCTCAGTACTTGCATTAT 216
 Db 2 TATACGGAAATACCTTTATCTCTGCTGATGCGAACACTTCTTTACTGACTTGCATTAT 61
 QY 217 ATTCTCCAGTAACCGGCTGCGAGGAAACAAGAACTGTCTGCCATAATCGGTAAATCTT 276
 Db 62 ATGCTCCGTGTGATTTCTGCTGGAAACAACAAGAACCGTGTGCCATAATCGGTAAATCTT 121
 QY 277 CTTGAGCATATAATTTAAATTCATCTGATGTTAAATGCGGATAGGAAATTTCTTCCAG 336
 Db 122 TTAGAGCATATAATTTCAAATTTTCAATTTGATTTAAACGCGGATAGGAGTTCTTCCAG 181
 QY 337 AAGACTGCCCCACATCGTGTATTTTACAAATGTCAGGAAGGTTGACATCATGTTTCATCAT 396
 Db 182 AAGACTGCCCCACATCGTGTATTTTACAAATGTCAGGAAGGTTGATCTCATGTACACCAC 241
 QY 397 TCAGCATGATGAATCAAAAACATCTGTTGAGGTTTCAAAAATCCAAAATCAAGAAAGAA 456
 Db 242 TCAGGCTGATGATCAAAAACATCTGTTGAGGTTTCAAAAATCCAAAATCAAGAAAGAA 301
 QY 457 CTTGATAGGTTGATCAATTTTTCAGAGATGTTTACTTATGATGATTTTAAAGAGAGTTTGGAG 516
 Db 302 CTTGATAGGTTGATCAATTTTTCAGAGATGTTTACTTATGATGATTTTAAAGAGAGTTTGGAG 361
 QY 517 AGCTTGGATTAACCTGCGGTATGATCTGAATGTTGATTTTCTAGATGTTCCATGACAGCAA 576
 Db 362 AGCTTGGATTAACCTGCGGTATGATCTGAATGTTGATTTTACTAGATGTTCCATGCGGACAA 421
 QY 577 AGCACAATTTTCATCGTTTTCAGAAAATTTCAATCTTAAATAACAATTCATGTCGCAAAAGTAGG 636
 Db 422 AGTACATTTTCATCGTTTTCAGAAAATTTCAATCTTAAATAACAATTCATGTCGCAAAAGTAGG 481
 QY 637 CTCAGAGAAATTTTCTCTCAAAAGATTAATCTTATTCAGGCGGTTTCTTCTGCTGAGTTG 696
 Db 482 CTCAGAGAAATTTTCTCTCAAAAGATTAATCTTATTCAGGCGGTTTCTTCTGCTGAGTTG 541
 QY 697 ACAGAGCAAGTTTCTCTGACCTTTCTGCTAGCAATATCAGATGCGGAGATATAGGATT 756
 Db 542 ACAGAGCAAGTTTCTCTGACCTTTCTGCTAGCAATATCAGATGCGGAGATATAGGATT 601
 QY 757 TCAATCTACGGAAGGAAACAGAGTGAATGGGACCAACTTGCAGTTGGATAGTGAACAAT 816
 Db 602 TCAATCTACGGAAGGAAACAGAGTGAATGGGACCAACTTGCAGTTGGATAGTGAACAAT 661
 QY 817 GAATTCACAGTGGAAATGTTGCTGCGGTGTTAGATTCACCGTTTATATATATATGTTAC 876
 Db 662 GAATTCACAGTGGAAATGTTGCTGCGGTGTTAGATTCACCGTTTATATATATATGTTAC 721
 QY 877 AAGGAATGGGTATCGTTACATCATCTTCCAAATCTTCTTGACAAATTTTCTGTTCTCTT 936
 Db 722 AAGGAATGGGTATCGTTACATCATCTTCCAAATCTTCTTGACAAATTTTCTGTTCTCTT 781
 QY 937 TTTGAGGTTTACTATTGATCCAGCTTTCACACCCACAGCTCCATGTCTCTCT 986
 Db 782 TTTGAGGTTTACTATTGATCCAGCTTTCATCCGAGCTTTCATCTCTCTCTCTCTCTCTCT 831

RESULT 4
 CB643461
 LOCUS
 DEFINITION OSJNEB04C17. f OSJNEB Oryza sativa (japonica cultivar-group) cDNA
 clone OSJNEB04C17 5', mRNA sequence.
 CB643461
 CB643461.1 GI:29638452
 EST

```

KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 798)
AUTHORS Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
JOURNAL Unpublished
COMMENT Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 04 row: C column: 17
Seq primer: gta aaa cga cgg cca gtc.
FEATURES
source
1..798
Location/Qualifiers
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSUNEB04C17"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSUNEB"
/note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
XhoI; 24 hrs after inoculation with Rice Blast (Che 86061)"
BASE COUNT 242 a 154 c 166 g 236 t
ORIGIN
Query Match 34.1%; Score 651.4; DB 14; Length 798;
Best Local Similarity 88.6%; Pred. No. 2.3e-54;
Matches 706; Conservative 0; Mismatches 91; Indels 0; Gaps 0;
QY 157 TGTACGGAGCATTTATCTCTGGCTGATGCTACAACTCTCTCAGTCTGCTAT 216
DB 2 TATACGGAATACCTTATTAATGTTCTGATGCGAACCTCTTTACTGACTTGCATTAT 61
QY 217 ATTCTCCGAGTAACGGCTGCGAGGAACAACAAGAACTGTCTGCATAATCGGTTAAATCTT 276
DB 62 ATGCTCCGGTGATTTCTCTGGAACAACAAGAACCGTGTGCCATAACCGGTTAAATCTT 121
QY 277 CTTGAGCATAGTTTAAATTCATCTGATGTTAAATGCGGATPAGGAATTTCTGCCGAG 336
DB 122 TTAGAGCATAAATTCAAATTTTCATTTGATGTTAAACGCGATPAGGAGTTCTTCCCCAG 181
QY 337 AAGACTGCCCCATCTGCTGATTTTACATGTCAGGAAGTTTGACATCATCTTCATCAT 396
DB 182 AAGACTGCACCATCTGCTGATTTTACATGTCAGGAAGTTTGATCTATCATGACCCAC 241
QY 397 TCAGCATGATGAATCAAAAACATCTGTTGAGGTTTCATAAAATCAAACTAAGAAAAGAA 456
DB 242 TCAGCGTGATGATCAATCAAAAACATCTGCTGAGATTATTAATCCAACTAAGGAAGAA 301
QY 457 CTTGATGAGGTGGTCAATTTTCAGAGATGTTATATGACTTTTAAAGAGGTTTTGAG 516
DB 302 CTTGATGAGGTGTCATTTTCAGAGACGGCACTTATATGACTCTCTGAAGGAGGTTTTGAG 361
QY 517 ACCTTGGACTTAACCTGGGTATCATCTGAATGTTGATTTTCTAGATGTCATGACACAAA 576
DB 362 AGCTTGGACTTAACCTGGGTATGACTTGAATGTTGATTTTACTAGATGTCATGCGGACAAA 421

```

```

QY 577 AGCATTTCATCGTTTTCACAAATTCCTAAATCTAAATCAATCCATGTGGCCAAAGTAGG 636
DB 422 AGTACATTTCATCGTTTTCACAAATTCCTAAATCTAAATCCATGTGGCCAAAGTAGG 481
QY 637 CTCAGAGAAATTTTCCTCACAACAGATATCTTATTCAAGCCGCTTTCTTGTCTGAGTTG 696
DB 482 CTCAGAGAGATTTTCCTCACAACAGAGCAATCTTATCAAGGTCGTTTCTTGTCCGAGCTA 541
QY 697 ACAAGCAAGTTTCTCTGACCTTCTCTAGCAATATCAGATGGCAGAAATATAGGATT 756
DB 542 ACAAGCAAGTTTCTCTGACCTTCTCTAGCAATATCAGATGGCAGAAATATAGGATT 601
QY 757 TCAATCTACGGAAGGAAACAGAGTGAATGGGACCAACTTGGCAAGTTGATAGTGAACAAT 816
DB 602 TCAATCTATGGAGGAGGAAAGTGAATGGGACCAACTTGGCAAGTTGATAGTGAACAAT 661
QY 817 GAATGTCACAGTGGAAATGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 876
DB 662 GAATGTCACAGTGGAAATGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 721
QY 877 AAGAAATGGGTATCGTTTACATCATCTTCAAAATCTTCTTGACAACTTTTCTTCTCTCT 936
DB 722 AAGAAATGGGTATGTTTACATCATCTTCAAACTCTTCTTGACAACTTTTCTTCTCTCT 781
QY 937 TTTGAGGTTACTATTGA 953
DB 782 GTTGAGGTCACCATTTGA 798
RESULT 5
AW017903 624 bp mRNA linear EST 13-SEP-1999
LOCUS 614020B05.y1 614 - root cDNA library from Walbot Lab Zea mays cDNA,
DEFINITION mRNA sequence.
ACCESSION AW017903
VERSION AW017903.1 GI:5871432
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 624)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL Unpublished
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 614020 row: B column: 05.
FEATURES
Location/Qualifiers
1..624
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="W23"
/db_xref="taxon:4577"
/tissue_type="root"
/dev_stage="3-4 days old"
/lab_host="X10LR"
/clone_lib="614 - root cDNA library from Walbot Lab"
/note="Organ: root; Vector: pBluescriptII SK+; Site 1:
EcoRI; Site 2: XhoI; 3-4 days old root tissue from Walbot
Lab (LM)"
BASE COUNT 164 a 149 c 133 g 176 t 2 others
ORIGIN
Query Match 31.7%; Score 606; DB 9; Length 624;

```

Best Local Similarity 99.7%; Pred. No. 6.3e-50;
Matches 606; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 972 GCTCCATGCTTCTCCTGAACGAGGTGTAGGGTTGGACCTGGTGTGATGATGAAGTAAACC 1031
Db 17 GCTCCATGCTTCTCCTGAACGAGGTGTAGGGTTGGACCTGGTGTGATGATGAAGTAAACC 76

QY 1032 AGAAGGGCTCCAAACAAACACATGCGCCACACCTGAACAGTGGACCAATGTGTCAACCC 1091
Db 77 AGAAGGGCTCCAAACAAACACATGCGCCACACCTGAACAGTGGACCAATGTGTCAACCC 136

QY 1092 TGCATTTTCATATTAATGCGTACTACTGCTATGCTATTAATTAATCAACCTAAACAGTGGC 1151
Db 137 TGCATTTTCATATTAATGCGTACTACTGCTATGCTATTAATTAATCAACCTAAACAGTGGC 196

QY 1152 TGAGTCAAAAGGGAATGACACATATCAAAATTCGGTCCACATGCTGGAGAGCTGGAGATGT 1211
Db 197 TGAGTCAAAAGGGAATGACACATATCAAAATTCGGTCCACATGCTGGAGAGCTGGAGATGT 256

QY 1212 TGATCACTTGGGAGGACATTTCTCTCTGTGCACACATATCACATGGAATTAATCTAAG 1271
Db 257 TGATCACTTGGGAGGACATTTCTCTCTGTGCACACATATCACATGGAATTAATCTAAG 316

QY 1272 GAAGTCTCTGTGCTTCAGTACTGTACTATCTTGTGTGATGCTGGCGATGCC 1331
Db 317 GAAGTCTCTGTGCTTCAGTACTGTACTATCTTGTGTGATGCTGGCGATGCC 376

QY 1332 ATTGAGCAACAACCTCTTATTTCTGACTATCATCGCAACCCCTTTTCCACGTTCTTCCA 1391
Db 377 ATTGAGCAACAACCTCTTATTTCTGACTATCATCGCAACCCCTTTTCCACGTTCTTCCA 436

QY 1392 ACGAGTCTGAATGTCTCATATCTACGATGACCTTTTGCATTAATCACTGACAAAAGA 1451
Db 437 ACGAGTCTGAATGTCTCATATCTACGATGACCTTTTGCATTAATCACTGACAAAAGA 496

QY 1452 ACCATTGGTGGAGAATACAGCAATTCGCTGCTGTGGAAGCTCAGTTCTTTGTGATTT 1511
Db 497 ACCATTGGTGGAGAATACAGCAATTCGCTGCTGTGGAAGCTCAGTTCTTTGTGATTT 556

QY 1512 ATGCGAAATGCGAGAACTCTGTTTACCAATCTGGGTTTTCATGCTCTCAAGCGCA 1571
Db 557 ATGCGAAATGCGAGAACTCTGTTTACCAATCTGGGTTTTCATGCTCTCAAGCGCA 616

QY 1572 CTGGATTG 1579
Db 617 CTGGATTG 624

RESULT 6
CB660024
LOCUS
DEFINITION OSJNE01H03.f OSJNE Oryza sativa (japonica cultivar-group) cDNA
ACCESSION
VERSION CB660024
KEYWORDS CB660024.1 GI:29663749
SOURCE
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzoideae; Oryza.
1 (bases 1 to 877)
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
Unpublished
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967

Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aac cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 01 row: H column: 03
Seq primer: gta aac cga cgg cca gtc.
Location/Qualifiers
1. 877
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNE01H03"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEd"
/note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
XhoI; 24 hrs after inoculation with Rice Blast (C9240-1)"
BASE COUNT 253 a 181 c 176 g 267 t
ORIGIN
Query Match 30.9%; Score 590.8; DB 14; Length 877;
Best Local Similarity 79.7%; Pred. No. 1.5e-48;
Matches 697; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

QY 655 AAACAAGATATCTTATTCAGGCGGTTTCTTCTGCTGAGTTGACAAAGCAAGTTTCTCT 714
Db 4 ATACAGGCAACCTTATTCAGGCGGATTTCTTGTCTGAATTGACAAAAGAGTATTTCT 63

QY 715 GACCTTTCTGTAGCAATATACAGATGCGAGAATATAGGATTTCAATCTACGAAAGGAAA 774
Db 64 GATCTTGAGCAAGTAATATACAGATGCTGATATAGAAATCTATCTATGGGAGAAAG 123

QY 775 CAGACTGAATGGGACCAACTTGCAGTTGGATGAGTAAAGTAAATGCAAGTGGCAAT 834
Db 124 AAAAGTGAATGGGATCAGATGCAAGTGGATAGTAAATGAAATGTAATGTAACGAGAA 183

QY 835 GTTGTCTGGCTGGTTTCAGATTCACGCTTATATATATGTAAGAAATGGTATCGTT 894
Db 184 GTTGTCTGGTTTAAATTCAGATTCCTCGGATATACAATGATATACAGGAGATGGGAA 243

QY 895 ACATCATTCACAAATCTTCTTGACAAATTTTGGTCTCTCTTTTGGAGTTACTATTGAT 954
Db 244 ATTCTTTCCAGAACTCTTTCAGAAATTTTCTGCTCTCTTTTGAAGTAACTGTTGAT 303

QY 955 CCAGCTTTCACACCCACAGCTCCATGCTTCTGTAAGAGAGGTTGTAGGTTGGACCTGGTT 1014
Db 304 CCGCTTTCACATCTCAGCTCCATGTTTCTTGCAACAGGTCGTTGGGCTGGATTTAGTG 363

QY 1015 GATGATGAAGTAAACCAAGGCGTCCAAAGAGCATGCCACACCTGGAACAGTGG 1074
Db 364 GATGATGAAGCAAAACCAAGAGAGAGCGCCCAACAAACATGATGATGATGATGATG 423

QY 1075 ACCAATGTTTCAACCCCTGCATTTTTCATATATGCTACTCTGCTACTGCTAACTATTTC 1134
Db 424 ACTAATGTTTCAATCCAGCATATGCTATATATGTTGTTGTTGTTGTTGTTGTTGTTG 483

QY 1135 ACCCTAAACAAGCTCGTGAAGTCAAGGGAATGACCATATCAATTCGTTCCACATGCT 1194
Db 484 ACGTGAACAAGCTTCGTTGAGTCCAAAGGATGATGACAAATCAAACTTCGTCACACTGT 543

QY 1195 GGAGAGCTGGAGATGTTGATCATCTGCGAGGACATTTCTTCTGTCACACATATCA 1254
Db 544 GGGAGGCTGGAGATGATGATCATCTTGTGCGAGCATTTCTTCTCTCTCTCTCTCTCT 603

QY 1255 CATGGAATTAATCTAAGGAAGTCTCTGTGCTTCTAGTACTTCTACTATCTCTGTCAGATT 1314
Db 604 CAGGGGTTAATTAAGAAAGTCCCTGCTCTCCAGTATCTGATATACCTAGCTCAGATT 663

QY 1315 GGTCTGGGATGTCCTCCCATTTGAGCAACAACTCTTTATTTCTTGTGACTATCATCGCAACCT 1374

Db 664 GGTCTTGCCATGCTCTTGTAGCAACAACTCATTTATTTGATTATATCACCGAACCT 723
 QY 1375 TTTCACAGCTTCTTCCACAGAGCTGTAAGTCTCATTTATCTACGATGACCTTTGCAA 1434
 Db 724 TTCCCAACATTTTCTCTAGAGGCTTACAGTTCTCTATCAACGATGACCTTTGCAA 783
 QY 1435 ATTACCTGACAAAAGAACCAATTGTTGGAAGATACAGATTGCTGCTGCTGTTGGAAG 1494
 Db 784 ATTACCTGACAAAAGAACCTTTGTTGAAGATATAGCATCGCTTCTGCTGTTGGAAG 843
 QY 1495 CTCAGTCTTGTGATTATTCGAAATTTGCGAGGA 1528
 Db 844 CTAAGTTCATGCGACCTATGTGAATTTGCTAGGA 877

RESULT 7
 AW671750
 LOCUS LGL_351_F02.bi_A002 Light Grown 1 (LGI) Sorghum bicolor cDNA, mRNA
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Sorghum bicolor (sorghum)
 Sorghum bicolor
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Sorghum.
 1 (bases 1 to 634)
 Cordonnier-Pratt M.-M., Gingle A., Marsala, C. and Pratt, L.H.
 An EST database from Sorghum: light-grown seedlings
 Unpublished
 Contact: Cordonnier-Pratt MM
 Laboratory for Genomics and Bioinformatics
 The University of Georgia, Department of Plant Biology
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
 Tel: 706 542 1860
 Fax: 706 583 0210
 Email: mmpratt@uga.edu
 Sequences have been trimmed to exclude PolyA, vector and regions
 below Phred quality 16. The threshold for highest quality sequence
 is 20.
 Seq primer: JEN REV
 High quality sequence stop: 516
 POLYA=No.

FEATURES
 source
 1..634
 /organism="Sorghum bicolor"
 /mol_type="mRNA"
 /db_xref="taxon:4558"
 /clone_lib="Light Grown 1 (LGI)"
 /note="Organ: 10- to 14-day-old light-grown (greenhouse)
 seedlings; Vector: Lambda Zap; Site 1: XhoI; Site 2: EcoRI
 ; The library was made from poly-A RNA in the cloning
 vector lambda ZAP II. Clones to be sequenced were
 prepared by mass excision."
 190 a . 137 c 126 g 181 t

BASE COUNT
 ORIGIN
 Query Match 30.2%; Score 576.8; DB 9; Length 634;
 Best Local Similarity 96.0%; Pred. No. 4.1e-47;
 Matches 603; Conservative 0; Mismatches 22; Indels 3; Gaps 1;

QY 523 GACTTAACCTGGGTATGATCTGAATCTGATTGCTAGATGTCCTAGCAGACAAAAGCACA 582
 Db 7 GCGTTAACTGGGTATGATCTGAATCTGATTGCTAGATGTCCTAGCAGACAAAAGCACA 66
 QY 583 TTTCATCGTTTGCACAAATTCATCTAAATAATCCATGTCGGCCAAAGTAGGCTCAGA 642
 Db 67 TTTCATCGTTTGCACAAATTCATCTAAATAATCCATGTCGGCCAAAGTAGGCTCAGA 126
 QY 643 GAAATTTCTCTCAACACAGATATCTTATTCAGGCGGTTTCTTGTGCTGAGTTGACAAAG 702

Db 127 GAAATTTCTCTCAACACAGACAAATCTCATTTCAAGGCGGTTTCTTGGCGAGTTGACAAAG 186
 QY 703 CAAATTTCTCTGACCTTTCTGTAGCAAAATATCAGATGCGAGATATAGGATTTCAATC 762
 Db 187 CAAATTTCTCTGACCTTTCTGTAGCAAAATATCAGATGCGAGATATAGGATTTCAATC 246
 QY 763 TACGAGAGGAACACAGATGGAATGGACCAACTTCGCAAGTTGGAATAGTGAACAATGAATTG 822
 Db 247 TATGGGAGGAACCAAGATGGAATGGACCAACTTCGCAAGTTGGAATAGTGAACAATGAATTG 306
 QY 823 CACAGTGGAAATGTTGCTGCTGCTGCTCAGATTCACGCTTATATATATGTTACAAAGGAA 882
 Db 307 CACATGGAATGTTGCTGCTGCTGCTCAGATTCACGCTTATATATGTTACAAAGGAA 366
 QY 883 ATGGGTATCGTTATCATCATTCCTCAAAATCTCTTTCGACCAACATTTGCTCTTTTTCAG 942
 Db 367 ATGGGTATGTTATCGTCAATTCCTCAAAATCTCTTTCGACCAACATTTTCTCTCTTTTTCAG 426
 QY 943 GTTACTATTGATTCAGCTTCACACCCACGCTCCATGCTCTTCTGAGCAGGTTGTAGGG 1002
 Db 427 GTTACTATTGATTCAGCTTCACACCCGAGCTCCATGCTCTTCTGAGCAGGTTGTAGGG 486
 QY 1003 TTGGACCTGGTTGATGATGAAGTAACACCAAGAGGCGTCCAAACAAAG--CACATGCC 1059
 Db 487 CTGACCTGGTTGATGATGAAGTAACACCAAGAGGCGTCCAAACAAAGCACCATGCT 546
 QY 1060 ACACCTGAACAGTGGACCAATGTTTCAACCTGCTTTCATATTTATGCTACTACTGTC 1119
 Db 547 ACACCTGAACAGTGGACCAATGTTTCAACCTGCTTTCATATTTATGCTACTACTGTC 606
 QY 1120 TATGCTAACTTATTCACCTTAAACAAGC 1147
 Db 607 TATGCTAACTTATTCACCTTAAACAAGC 634

RESULT 8
 CB668055
 LOCUS OSJNEd15L08.f OSJNEd Oryza sativa (japonica cultivar-group) cDNA
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Oryza sativa (japonica cultivar-group)
 Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Eriactoidae; Oryzaceae; Oryza.

1 (bases 1 to 845)
 Jantasuriyarat, C., Lu, G., Gowda, M., Hatfield, J., Zhou, B., Mazur, E.,
 Kudrna, D., Dean, R., Soderlund, C., Wing, R. and Wang, G.
 Large-scale identification of ESTs involved in the interaction
 between rice and Magnaporthe grisea
 Unpublished
 Contact: Rod Wing
 Arizona Genomics Institute
 University of Arizona
 Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
 85721-0088, USA
 Tel: 520 626 3967
 Fax: 520 621 9288
 Email: http://genome.arizona.edu
 PCR Primers
 FORWARD: gta aaa cga cgg cca gtc
 BACKWARD: gga aac agc tat gac cat g
 Plate: 15 row: L column: 08
 Seq primer: gta aaa cga cgg cca gtc.

FEATURES
 source
 1..845
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"


```

QY 312 TCGGGATAGGAATTTCTTGGCCAGAGAGACTGCCCCACATCGTGATTTTACAATGTCAG 371
Db 310 -----GGAGTTTCTTGGCCAGAGAGACTGCCCCACATCGTGATTTTACAATGTCAG 360
QY 372 GAAGGTTGACACTCATGTTTCATCTTCCAGCATGCAATCAAAACATCTGTTGAGGTT 431
Db 361 GAAGGTTGATATCATGTTACACCATCAGCGTGATGAATCAAAACATCTGTTGAGAT 420
QY 432 CATAAATCAAACTAAGAAAAGAACCTGATGAGTGGTCAATTTTCAGAGATGGTACTTA 491
Db 421 TATAAATCAAACTAAGAAAAGAACCTGATGAGTGGTCAATTTTCAGAGAGCGCACTTA 480
QY 492 TATGACTTTAAAGAGGTTTTTGAGAGCTTGAATTAACCTGGGTATGATCTGAATGTTGA 551
Db 481 TATGACTCTGAAGAGGTTTTTTGAGAGCTTGAATTAACCTGGGTATGATCTGAATGTTGA 540
QY 552 TTTGCTAGATGTCATGTCAGACAAAAGCACATTTTCATCGTTTTTGACAAATTCAACTCAA 611
Db 541 TTTTCTAGATGTCATGTCAGACAAAAGCACATTTTCATCGTTTTTGACAAATTCAACTCAA 600
QY 612 ATACATCCATGTCGGCCAAAGTAGCTCAGAGAAATTTTCTCAAAACAGAGATAATCTTAT 671
Db 601 GTACATCGTGTGGCCAAAGTAGCTCAGAGAGATTTTCTCAAAACAGAGATAATCTTAT 660
QY 672 TCAAGCCGTTTTCTTGTGAGTTGACAAAGCAAGTTTTCTCTGACCTTCTGCTAGCAA 731
Db 661 CCAAGGTCGTTTTCTTGTGCGAGCTTCAAAAGCAAGTTTTCTCTGACCTTACAGCAAGTAA 720
QY 732 ATATCAGATGGCAGATATAGGATTTCAATCTACGG 767
Db 721 ATACAGATGGCTGAATATAGGATTTCAATCTATGG 756

```

RESULT 10

```

BM359276
LOCUS BM359276
DEFINITION GA_Ea0018J03r Gossypium arboreum 7-10 dpa fiber library Gossypium
          arboreum cDNA clone GA_Ea0018J03r, mRNA sequence.
ACCESSION BM359276
VERSION   BM359276.1
KEYWORDS  EST.
SOURCE    Gossypium arboreum
ORGANISM  Gossypium arboreum

```

REFERENCE

```

AUTHORS  Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
          ,D., Wood,T.C., Leslie,A. and Wilkins,T.A.
TITLE     An integrated analysis of the genetics, development, and evolution
          of the cotton fiber
JOURNAL   Unpublished
COMMENT   Contact: Wing RA
          Clemson University Genomics Institute
          Clemson University
          100 Jordan Hall, Clemson, SC 29634, USA
          Tel: 864 656 7288
          Fax: 864 656 4293
          Email: rwing@clemson.edu
          Total High Quality bases = 590
          Seq primer: TAATACGACTCATATAGG
          High quality sequence stop: 728.

```

FEATURES

```

source
1..872
/organism="Gossypium arboreum"
/mol_type="mRNA"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ea0018J03r"
/tissue_type="Fibers isolated from bolls harvested 7-10
          dpa"

```

RESULT 11
BU973129

```

/lab_host="E. coli"
/clone_lib="Gossypium arboreum 7-10 dpa fiber library"
/ncbi_vector="pBK-CMV; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 257 a 174 c 164 g 271 t 6 others
ORIGIN

```

```

Query Match 27.1%; Score 517; DB 12; Length 872;
Best Local Similarity 75.4%; Pred. No. 2e-41;
Matches 651; Conservative 0; Mismatches 211; Indels 1; Gaps 1;

QY 574 AAAAGCAGATTTTCATCGTTTTTGACAAATTTCAATCTTAAATAACAATCCATGTCGCAAGT 633
Db 2 AAGAGCACTTTTCATCGTTTTTGACAAATTTCAATCTTAAATAACAATCCATGTCGCAAGT 61
QY 634 AGGCTCAGAGAAATTTTCTCAAAACAAGATATCTTATCAAGGCGGTTTTCTTGTCTGAG 693
Db 62 AGACTTAGAGAGATCTTTTAAAGCAGGACAATCTTATCCAAGGAGGTTTTCTGCGAGAA 121
QY 694 TTGACAAAGCAAGTTTTCTCTGACCTTTCTGCTAGCAAAATATCAGATGGCAGAAATATAGG 753
Db 122 GTTACAAAGCAAGTCTATCAGATCTTGAAACAAGCAAAATATCAGATGGCAGAGTACAGG 181
QY 754 ATTTCAATCTACGGAAGGAAACAGAGTGAATGGGACCAACTTTGCAAGTTGGATAGTGAAC 813
Db 182 GTGTCATATATGAAGGAAACAAAGTGAATGGGACCAAGTTGGCAAGTTGGTTATTATAC 241
QY 814 AATGAATTCACAGTGGAAATTTTGTCTGGCTGGTTCCAGATTCCACGCTTATATAATGTG 873
Db 242 AATGAATTTATAGTGAGACTTACTGTTGGTTAATCCAGTTTACCAAGATTCTATAATGTG 301
QY 874 TACAAGAAATGGGTATCGTTACATCATTCGAAATCTTCTTGACAAATTTCTGTCCTTCT 933
Db 302 TACAAGCAATGGGAATTTGTAATAATCTTTTCAAGAAATTTCTGGAATATTTCTGACAT 361
QY 934 CTTTTCGAGTTACTATTGATCCAGCTTCACCCACAGCTCCATGCTCTTCTGAGCAG 993
Db 362 CTATTGAGTTACAGTTGATCCAAATTTCTCATCTCAACTACACGTTGTTCTTAAAAATG 421
QY 994 GTTGTAGGTTGGACCTGGTTGATGATGAAGTAAACAGAGGCGTCCCAACAAAGCAC 1053
Db 422 GTGGTTGGCTTTGATTAGTTGATGAGAGTAAACAGAGAGCGCTCTTACAAAGCAC 481
QY 1054 ATGCCCAACCTGAAACAGTGGACCAATGTGTTCAACCTCGCATTTTCATATTATGCGTAC 1113
Db 482 ATGCCCAACTCTGCTGAATGGACTTAATGAAATTTAATCTGCTATTTCTTATTATGCTTAC 541
QY 1114 TACTGCTATGCTAACTTATTCACCTTAAACAGCTGGTGGTCAAGGGGAATGACCACT 1173
Db 542 TACTTTTATGCAAACTTTTATACACTCAACAGCTTCGTGAGTCAAAAGGAATGCAGACA 601
QY 1174 ATCAAAATTCGTCACATGCTGGAGAGCTGGAGATGTTGATCACTTGGCAGCGACATTT 1233
Db 602 ATAAACTCCGACCTCTACTGNGGAGGCTGGTATATTGACCATNTAGCTGCTGCTTTC 661
QY 1234 CTTCTGTGCAACAATATCAATGGAATTAATCTAAGGAAGTCTCTGTCGCTTTCAGTAC 1293
Db 662 CTTCTGTG-CACANTATATCTCACGGGATTAATCTNCGAAATCCCTGTTTTCAGTAC 720
QY 1294 TTGTACTATCTTGTGTCAGATTGGTCTGGCGATGTCCTCATTTGAGCAACAACCTCTATT 1353
Db 721 TTGTATTACCTCGCTCAGATCGGATNGGCAATGTCCTCTGAGCAATAATATCCCTTTTC 780
QY 1354 CTTGACTATCATCGCAACCTTTTCCAAAGTTCTTCCAAAGGCTGCTGAATGTCTCATTA 1413
Db 781 CTGACTATCATCGCAACCTTTTCTTCTTCTTTCAGGGGCGCTTAATGTTTCTTCTT 840
QY 1414 TCTACGATGACCTTTTGCAAAT 1436
Db 841 TCATCTGATGACCTTAACTT 863

```

LOCUS BU973129 647 bp mRNA linear EST 22-OCT-2002
 DEFINITION HB23N17: BC Hordeum vulgare subsp. vulgare cDNA clone HB23N17
 5-PRIME, mRNA sequence.
 ACCESSION BU973129
 VERSION BU973129.1 GI:24223322
 SOURCE EST.
 ORGANISM Hordeum vulgare subsp. vulgare
 Hordeum vulgare subsp. vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
 ; Triticeae; Hordeum.
 1 (bases 1 to 647)
 REFERENCE Radchuk,V., Zhang,H., Weschke,W., Potokina,E. and Wobus,U.
 AUTHORS Barley ESTs from developing seeds
 TITLE Unpublished
 JOURNAL
 COMMENT Contact: Stein Nils
 Molecular Markers Group, Department Genbank
 Institute of Plant Genetics and Crop Plant Research (IPK)
 Corrensstr. 3, 06466, Gatersleben, Germany
 Tel: 039482-5522
 Fax: 039482-5595
 Email: stein@ipk-gatersleben.de
 Insert Length: 647 Std Error: 0.00
 Plate: 23 row: N column: 17
 Seq primer: M13rev.

FEATURES

source

1. .647
 Location/Qualifiers
 /organism="Hordeum vulgare subsp. vulgare"
 /mol_type="mRNA"
 /cultivar="barke"
 /db_xref="GABI:237971"
 /db_xref="taxon:112509"
 /clone="HB23N17"
 /tissue_type="developing caryopsis"
 /dev_stage="8-15 DAP (days after pollination)"
 /lab_host="XL10-Gold"
 /clone_lib="BC"
 /notes="Vector: pBluescript SK+; Site 1: EcoRI (5'-end of cDNA); Site 2: XhoI (3'-end of cDNA); developing caryopsis
 , 8-15 DAP (days after pollination) Due to a cloning
 artefact caused by the kit, in most cases the EcoRI site
 is NOT present, as well as the EcoRI adapter used for
 cloning. To excise the insert, restriction sites upstream
 EcoRI should be used (e.g. BamHI, SalI, PstI). NOTE: Also
 due to the cloning system used Blue/white selection for
 recombinants is not 100% reliable."

BASE COUNT 176 a 136 g 187 t
 ORIGIN

Query Match 26.6%; Score 508.2; DB 13; Length 647;
 Best Local Similarity 87.1%; Pred. No. 1.7e-40;
 Matches 558; Conservative 0; Mismatches 83; Indels 0; Gaps 0;
 QY 1073 GGACCAATGTGTTCAACCCCTGCATTTTCATATTATGCGTACTACTGCTACTGCTAACTTAT 1132
 Db 1 GGACCAATGTGTTCAACCCCTGCATTTTCATATTATGCGTACTACTGCTACTGCTAACTTGT 60
 QY 1133 TCACCCTAACACAGCTCGGTGAGTCAAGGAGATGACCACTATCAAAATTCGCTCCACATG 1192
 Db 61 ACACACTGAACAAGCTGCGTGAAGTCAAGGGGATGAATATCAAAATTCGCTCCACATG 120
 QY 1193 CTGAGAGGCTGGAGATGTTGATCATTGTCAGCGACATTTCTTCTCTGTCAACAACATAT 1252
 Db 121 CTGAGAGGCTGGAGATGTTGATCATTGTCAGCGACATTTCTTCTTGTGACAGATATAT 180
 QY 1253 CACATGGAATTAATCTAAGGAAGTCTCCTGTGCTTCAAGTACTTGTACTATCTTGGTCAGA 1312
 Db 181 CACATGGAATCAATTTAAGGAAGTCTCCTGTGCTTCAAGTACTTGTACTATCTTGGTCAGA 240
 QY 1313 TTGGTCTGGCGATGTCGCCATTGAGGCAACACTCCTTATTTCTTGTACTATCATCGCAACC 1372
 Db 241 TCGGTCTAGCAATGTGCCCAATTAAGCAACAACCTCCTTGTCTTGTGATTACCATCGGAACC 300

QY 1373 CTTTTCACACGTTCTTCCACAGAGGTCTGAATGCTCAATTATCTACGATGACCCCTTGC 1432
 Db 301 CTTTTCCTATGTTTTCACAGAGACTGAATGCTCAATTATCTACGATGATCCATTCG 360
 QY 1433 AAATTCACCTGACAAAGAACCATTTGGTGAAGATACAGCATTCCTCTCGCTGCTGA 1492
 Db 361 AAATTCATCTGACAAAGAGCGCTTGGTCGAGGAATACAGCATTCGCCCTTCGCTGTGA 420
 QY 1493 AGCTCAGTCTCTTGATTTATGCGAAATTCGAGGAACCTCTGTTTACCAATCTGGGTTT 1552
 Db 421 AGCTCAGTCTCTTGATTTATGTAATTCGAAGAAATCTGTGTATCAATCAGGCTTT 480
 QY 1553 CACATGCTCTCAAGCGGCACCTGGATTGTAAGAACTACTTCAAAAGAGAGACTGCTGAA 1612
 Db 481 CACATGCTCTCAAGGCACATTTGGATTGGCAAGAACTACTACAAGAGAGGCCCTCAGGA 540
 QY 1613 ATGATATTACAGAACCAATGTACCGCACATCAGGTTCAATTTAGAGAGATGATCTGA 1672
 Db 541 ACGATATATCCAGAACGAATGTGCCACCACATCAGGATCGAATTCAGGACCTGATCTGA 600
 QY 1673 GAAATGAATGAAACTAGTGTACTCTGACAATGAGATCTTA 1713
 Db 601 GAGAGAAATGCGAGCTCGTTTACCTTGATACGTCATCTTA 641

RESULT 12

BI955474

LOCUS BI955474 676 bp mRNA linear EST 19-OCT-2001
 DEFINITION HVSMEM0023E15f Hordeum vulgare green seedling EST library
 HVCDNA0014 (Blumeria infected) Hordeum vulgare subsp. vulgare cDNA
 clone HVSMEM0023E15f, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

BI955474.1 GI:16301827
 EST
 Hordeum vulgare subsp. vulgare
 Hordeum vulgare subsp. vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
 ; Triticeae; Hordeum.
 1 (bases 1 to 676)

REFERENCE

AUTHORS

TITLE

Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Chin,A., Begum,D.,
 Frisch,D., Atkins,M., Yu,Y., Henry,D., Palmer,M., Rambo,T., Simmons
 J., Oates,R. and Main,D.
 Development of a genetically and physically anchored EST resource
 for barley genomics: Blumeria infected Morex (compatible) seedling
 cDNA library

JOURNAL

COMMENT

Unpublished
 Contact: Wing RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Total hg bases = 442
 Seq primer: AATTAACTCTCACTAAAGGG
 High quality sequence start: 23
 High quality sequence stop: 541.
 Location/Qualifiers
 1. .676
 /organism="Hordeum vulgare subsp. vulgare"
 /mol_type="mRNA"
 /cultivar="Morex"
 /db_xref="taxon:112509"
 /clone="HVSMEM0023E15f"
 /tissue_type="green seedling leaf"
 /lab_host="TJCI21"
 /clone_lib="Hordeum vulgare green seedling EST library
 HVCNA0014 (Blumeria infected)"
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
 XhoI; Morex (mla) plants were greenhouse grown in the R
 Wise lab at Iowa State University, Ames, IA; 7 day old
 green seedlings were infected with isolate 5874 of

FEATURES

source

Blumeria graminis f. sp. *hordei*, and leaves were harvested 24, 48 and 72 hr post-inoculation and snap frozen (Wise). In the TU Close lab at the University of California, Riverside, total RNA was prepared from each sample pool, equal quantities of all three RNA pools were combined, poly(A) RNA was purified from the mixture, one primary unamplified cDNA library was made, and 1 million pfu were in vivo excised to give pluescript SK(-) cDNA phagemids (Chin). Phagemids were plated and picked at the Clemson University Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA preparations, DNA sequencing and sequence analysis were performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main). The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see <http://www.genome.clemson.edu/projects/barley>. To order this clone see <http://www.genome.clemson.edu/orders> Also see Close TU, Wing R, Kleinof A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. *Barley Genetics Newsletter* 31:29-30. (<http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html>)

```

BASE COUNT      186 a 132 c 153 g 205 t
ORIGIN

Query Match      26.6%; Score 507.6; DB 12; Length 676;
Best Local Similarity 87.6%; Pred. No. 1.9e-40;
Matches 578; Conservative 0; Mismatches 79; Indels 3; Gaps 2;

QY 777 GAGTGAATGGGACCAACTTGCAGTTGGATAGTGAACATGAATTCACAGTGGAAATGT 836
Db 18 GAGTGAATGGGACCAACTTGCAGTTGGATAGTGAACATGAATTCACAGTGGAAATGT 77

QY 837 TGCTGGCTGGTTTCAGATTCACAGCTTATATATATGTGTACAAAGAAATGGGTATCGTTAC 896
Db 78 TGTTGGCTGATTCAGATTCACAGCTTATATATATGTGTACAAAGAAATGGGTATCGTTAC 137

QY 897 ATCATTCCAAAATCTTCTTGACAACATTTTCGTTCTCTTTTGGAGTTACTATTGATCC 956
Db 138 ATCATTCCAAAATCTTCTTGACAACATTTTCGTTCTCTTTTGGAGTTACTATTGATCC 197

QY 957 AGCTTCACACCCACAGCTCCATGCTTCTGAGCAGGTTGTAGGGTTGGACCTGGTTGA 1016
Db 198 AGCTTCGACCCGACGCTTCATGCTCTCTTAAAGCAGGTTGTAGGGTTGGATTTGGTTGA 257

QY 1017 TGATGAAGTAAACAGAAAGCGCTCCAAACAAAGCACATGCCACACCTGACACGTGGAC 1076
Db 258 TGATGAGAGTAAACCTGAAAGCGTCCAACTAAGCATATGCTACACCTGAAAGATGGAC 317

QY 1077 CAATGTGTTCAACCCCTGCAATTTTCATATATGCTACTACTGCTATGCTAACTTATTCAC 1136
Db 318 AAATGTCTTCAACCCCTGCAATTTTCATATATGCTACTACTGCTATGCTAACTTATTCAC 377

QY 1137 CCTAAACAGCTGGTGAGTAAAGGAATGACCATATCAAAATCCGTCACATGCTGG 1196
Db 378 ACTGAACAGCTGGTGAGTAAAGGGATGAATATCAAAATCCGTCACATGCTGG 437

QY 1197 AGAGCTGGAGATGTTGATCACTTCGCGAGCAGATTTCTCTCTGTCACAAATATACCA 1256
Db 438 TGAGCTGGAGATGTTGATCACTTCGCGAGCAGATTTCTCTCTGTCACAAATATATACCA 497

QY 1257 TGAATTAATCTAAGGAAGTCTCCTGTGCTTCAGTACTTGTACTATCTTGGTCAGATTGG 1316
Db 498 TGAATTAATCTAAGGAAGTCTCCTGTGCTTCAGTACTTGTACTATCTTGGTCAGATTGG 557

QY 1317 TCTGGGATGTCCTTATGAGCAACATCTCCTTATTTCTTGACTATATCGCAACCCCTTT 1376
Db 558 GGTAGCAATGTCCTTATGAGCAACATCTCCTTATTTCTTGACTATATCGCAACCCCTTT 617

QY 1377 TCCAACGCTTCTTCCAACGAGTCTGAATGCTCATTATCTAGGATGACCTTTCGAAT 1436
Db 618 CCTA--TGTTTTTCAACGAGGACTGAATG-GTTATTTGTCACGGATGATCCATTGCAAT 674

```

RESULT 13 BF268586

LOCUS 880 bp mRNA linear EST 07-MAR-2001
DEFINITION GA_EB0002D19f Gossypium arboreum 7-10 dpa fiber library Gossypium
ARBOREUM CDNA clone GA_EB0002D19f, mRNA sequence.
ACCESSION BF268586
VERSION BF268586.2 GI:13247099
KEYWORDS EST.
SOURCE Gossypium arboreum
ORGANISM Gossypium arboreum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 880)
Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
D., Wood,T.C., Leslie,A., and Wilkins,T.A.
An integrated analysis of the genetics, development, and evolution
of the cotton fiber
Unpublished
On Nov 17, 2000 this sequence version replaced gi:11199581.
Contact: Wing RA
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: TAATACGACTCACTATAGG
High quality sequence stop: 838.
Location/Qualifiers
1..880
/organism="Gossypium arboreum"
/mol_type="mRNA"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_EB0002D19f"
/tissue_type="Fibers isolated from bolls harvested 7-10
dpa"
/lab_host="E. coli"
/clone_lib="Gossypium arboreum 7-10 dpa fiber library"
/note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"

FEATURES source

```

BASE COUNT      249 a 176 c 176 g 278 t
ORIGIN

Query Match      25.9%; Score 494.6; DB 10; Length 880;
Best Local Similarity 74.6%; Pred. No. 2.9e-39;
Matches 634; Conservative 0; Mismatches 215; Indels 1; Gaps 1;

QY 778 AGTGAATGGGACCAACTTGCAGTTGGATAGTGAACATGAATTCACAGTGGAAATGT 837
Db 1 AGTGAATGGGACCAACTTGCAGTTGGATAGTGAACATGAATTCACAGTGGAAATGT 60

QY 838 GTCTGGCTGGTTTCAGATTCACAGCTTATATAATGTGTACAGAAATGGGTATCGTTACA 897
Db 61 GTTTGGTTAATCCAGTTACACAGATTTGTAATGTGTACAGCAAAATGGGAATGTAAA 120

QY 898 TCATTCCAAAATCTTCTTGACAAACATTTTCGTTCTCTCTTTTGGAGTTACTATTGATCCA 957
Db 121 TCTTTTCAGATAATCTTGGACAATGTTTTCATTCCACTATTTTGAAGTTACAGTTGATCCA 180

QY 958 GCTTCACACCCACAGCTCCATGCTCTCTGAGCAGGTTGTAGGGTTGGACCTGGTTGAT 1017
Db 181 AATTCTCATCTCAACTACAGCTGTTCTTAAATAATGGTGGTTGGCTTTGATTTAGTTGAT 240

QY 1018 GATGAAGATGAACAGAAAGCGTCCAAACAAAGCACATGCCACACCTGAAACAGTGGACC 1077
Db 241 GATGAGATGAACAGAGAGCGGCTCTTACAAGCACATGCCAACTCTCTGCTGAATGNACT 300

QY 1078 AATGTTTCAACCTCGATTTTCATATATATGCGTACTACTGCTATGCTTAATCTTATTCACC 1137

```

```
Db 301 AATGAATTTAATCCTGCATATTTCTTATTATGCTACTACTTTTATGCAAAACCTTTATACA 360
QY 1138 CTAACAAGCTGCTGAGTCAAGGAATGACACATATCAAAATCCGTCACATGCTGA 1197
Db 361 CTAACAAGCTTCTGAGTCAAAAGGAATGACACAATAAACTCCGACCTCAGTGGG 420
QY 1198 GAGGCTGGAGATTTGATCACCTTTGGCAGGACATTTCTTCTCTGTGTCACAACTATACAT 1257
Db 421 GAGGCTGGTATGATTCACCATTTAGCTGCTGCTTCTCTGTGCAACAATAATCTCAC 480
QY 1258 GGAATTAATCTAAGGAAGTCTCTGTGCTTCAGTACTGTGATATCTTGGTCAGATTGTT 1317
Db 481 GGGATTAATCTCCGGAATCCCTGTTTGGCAGTACTGTGATTAACCTCGCTCAGATCGGA 540
QY 1318 CTGGGATCTCCCATTTGAGCAACACTCTTATTTCTTGTGACTATCATCGCAACCTTTT 1377
Db 541 TTGGCCATGCTCTCTGAGCAATAATTCCTTTCTTGAGACTATCATCGCAACCAATTT 600
QY 1378 CCAACGTTCTTCCAAAGAGTCTGAATGTCTCATTTATCTACGGATGACCTTTTGCAAAAT 1437
Db 601 CTTTCATTTCTCAGCTGGCTAAATGTTCTTCTTCATCTGATGACCTTTTACAAAT 660
QY 1438 CACCTGACAAAAGAACATTTGGTGAAGAAATACAGCATTTGCTGCTTGGTGGGAAGCTC 1497
Db 661 CATTTCACGAAGGAAGCGCTTGTGGAAGAGTACAGTGTTCGACACAGGTTTGGGAAGCTC 720
QY 1498 AGTTCTTGTGATTTATGCGAAATTCGAGGAACTCTGTTTACCAATCTGGTTTTTCACAT 1557
Db 721 AGTGTCTGACATTTATGTGAGATTTGTAGAAATCTGTATACCAAGTCTGGTGTGTCAC 780
QY 1558 GCTCTCAAGCGGCACAT-GAATTTGTAAGAACTACTTCAAAGAGGACCTGCTGGAATGA 1616
Db 781 ATGTCAAGTTGATGCTGGGCTTGGCAACAATAATTTCTTAAGAGGTCGAGAGGATGA 840
QY 1617 TATTACAGA 1626
Db 841 CANTCAAAA 850

RESULT 14
BG365787
LOCUS
DEFINITION
913 bp mRNA linear EST 22-OCT-2001
HVSMEI0004D04f Hordeum vulgare 20 DAP spike EST library HVCNDA0010
(20 DAP) Hordeum vulgare subsp. vulgare cDNA clone HVSMEI0004D04f,
mRNA sequence.
ACCESSION
BG365787.2 GI:16323876
VERSION
EST.
KEYWORDS
Hordeum vulgare subsp. vulgare
SOURCE
Hordeum vulgare subsp. vulgare
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Hordeum.
REFERENCE
1 (bases 1 to 913)
Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Begum, D., Frisch, D., Yu
Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi, D.W., Fenton
R.D., Close, S.J., Oates, R. and Main, D.
Development of a genetically and physically anchored EST resource
for barley genomics: Morex 20 DAP spike cDNA library
Unpublished
On Mar 8, 2001 this sequence version replaced gi:13254886.
CONTACT
Contact: Wing RA
Clemson University
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: twing@clemson.edu
Total hg bases = 406
Seq primer: AATAACCTCCTAAGGG
High quality sequence start: 2
High quality sequence stop: 622.
Location/Qualifiers
```

```
source
1. 913
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Morex"
/db_xref="taxon:112509"
/clone="HVSMEI0004D04f"
/tissue_type="20 DAP spike"
/lab_host="SOLR"
/clone_lib="Hordeum vulgare 20 DAP spike EST library
HVCNDA0010 (20 DAP)"
/note="Vector: lambdaZAP; Site 1: EcoRI; Site 2: XhoI;
Plants were grown in the greenhouse at the University of
California, Riverside (Fenton, SJ Close, TJ Close). Whole
spikes with awns trimmed were collected at 20 DAP (Fenton
). Total RNA was prepared, poly(A) RNA was purified, one
primary unamplified cDNA library was made, and 1 million
pfu were in vivo excised to give pBluescript SK(-) cDNA
phagemids in the TJ Close lab at the University of
California, Riverside (Choi). Phagemids were plated and
picked at the Clemson University Genomics Institute (CUGI)
(Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA
preparations, DNA sequencing and sequence analysis were
performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates
, Rambo, Main). The sequence has been trimmed to remove
vector sequence and contains a minimum of 100 bases of
phred value 20 or above. For more details on library
preparation and sequence analysis see
http://www.genome.clemson.edu/projects/barley. To order
this clone see http://www.genome.clemson.edu/orders/Also
see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
Genetically and physically anchored EST resources for
Barley genomics. Barley Genetics Newsletter 31:29-30.
(http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"
BASE COUNT 274 a 189 c 187 g 262 t
ORIGIN
Query Match 25.7%; Score 490.6; DB 10; Length 913;
Best Local Similarity 84.3%; Pred. No. 6.9e-39;
Matches 575; Conservative 0; Mismatches 105; Indels 2; Gaps 2;
QY 12 GGTTCCTCATGGGAGAGAGGCTATAAATGACCCCTGTACTCCAAAACCTAACCCNA 71
Db 232 GGTCTCTCATGGGAGAGAGGCTATAAATGACCCCTGTACTCCAAAACCTAACCTAA 291
QY 72 CCGCTTCACTTATGCTGCTGAACCAAGTCAGAGCATGTTTTCACAACTGTTGATGGGT 131
Db 292 CCCATTCATTTCTGCTGCTGAACCAAAATCGGAGCATGTTTTCACAAATGGTCATGTT 351
QY 132 TATCATGTTTATGCGGATAAAGATTGTACGGAGAGCATTTATCTGTGGCTGTACTAC 191
Db 352 TGTCCAGGCTATGCGGATAAAGACTGTACGGAAAGGATTTATCTGTTGCTGTACTAC 411
QY 192 AACCTTCTCACTGACTGCTGATTAATATTTCTCCGAGTAACGGCTCAGGGAACACAGAAC 251
Db 412 AACCTTTTTCACCGACTTGCATTAATGTTCTCCGGTGACTGCGCGGGGAATACAGAAC 471
QY 252 TGTCTGCCATATCCGTTAAATCTTCTGAGCATAGTTTAAATTCATTCATCTGATTTAA 311
Db 472 TGTCTGTATACACCGGTTGAATCTTCTAGAGCATAGTTTCAAAATTCATCTGATTTAA 531
QY 312 TCGCGATAGGGAATTTCTTTGCCAGAGAAGACTCCCCACATCTGTGATTTTTTACAATGT 371
Db 532 CGCGGACAGGAGTTTCTTGGCAGAGACTGCACCACATCTGTGATTTTACAATGTTAG 591
QY 372 GAAGTTGACACTCATGTTTCATCTCAGCATGATGATGATGATGATGATGATGATGAT 431
Db 592 GAAAGTCGACACTCATGTTTCAACCTCAGCATGATGATGATGATGATGATGATGAT 651
QY 432 CATAAATCCCAACTAAGAAAGAACCTTCATGAGGTGGTTCATTTTCAGAGATGGTACTTA 491
Db 652 CATCAAGTCCAAACTTGGAAAGAACCTTCATGAGGTGGTTCATTTTCAGAGATGGTACTTA 711
QY 492 TATG-ACITTTAAAGAGAGGTTTTTTGAGAGCTTGAAGCTTAACTGGGTATGATCTGTAATGTTG 550
```



```

Db      712 TATGACCTTAAGGAGGTTTGGAGCTTGACCTTAACCTGGTATGACTGGAATGTTG 771
Qy      551 ATTGCTAGATGTCATGCGAGCAAAAGCACATTTTC-ATCGTTTGGACAAATTCATCTTA 609
Db      772 ATTGCTAAATGTCACGCTGACAAAAGTACTTTTCAATCGTTTGAACAAATCAAACCTT 831
Qy      610 AATACATCCATGTCGCGCAAGTAGGCTCAGAGAAATTTTCTCAACAAAGATATCTT 669
Db      832 AATACCATCATGCGCGCAAGTAGTTACGCGAAATTTCTCTAAACGGGACATCT 891
Qy      670 ATTCAAGGCGTTTCTTCTGCTG 691
Db      892 TATTAAGGCAATTTTGGCTG 913

```

```

RESULT 15
LOCUS   BG442858
DEFINITION GA_Ea0018J03f Gossypium arboreum 7-10 dpa fiber library Gossypium
ARBOREUM cDNA clone GA_Ea0018J03f, mRNA sequence.
ACCESSION BG442858
VERSION   BG442858.1
KEYWORDS EST.
SOURCE   EST.
ORGANISM Gossypium arboreum

```

```

REFERENCE 1 (bases 1 to 909)
AUTHORS  Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
          D., Wood,F.C., Leslie,A. and Wilkins,T.A.
TITLE    An integrated analysis of the genetics, development, and evolution
          of the cotton fiber
JOURNAL  Unpublished
COMMENT  Contact: Wing RA
          Clemson University Genomics Institute
          Clemson University
          100 Jordan Hall, Clemson, SC 29634, USA
          Tel: 864 656 7288
          Fax: 864 656 4293
          Email: rwing@clemson.edu
          Seq primer: TAATACGACTCACTATAGG
          High quality sequence stop: 714.
          Location/Qualifiers
            1..909
              /organism="Gossypium arboreum"
              /mol_type="mRNA"
              /strain="AKA"
              /cultivar="8400"
              /db_xref="taxon:29729"
              /clone="GA_Ea0018J03f"
              /tissue_type="fibers isolated from bolls harvested 7-10
              dpa"
              /lab_host="E. coli"
              /lab_lib="Gossypium arboreum 7-10 dpa fiber library"
              /note="Vector: pBK-CMV; Site 1: EcoRI; Site 2: XhoI"

```

FEATURES

```

source
1..909
/mol_type="mRNA"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ea0018J03f"
/tissue_type="fibers isolated from bolls harvested 7-10
dpa"
/lab_host="E. coli"
/lab_lib="Gossypium arboreum 7-10 dpa fiber library"
/note="Vector: pBK-CMV; Site 1: EcoRI; Site 2: XhoI"
BASE COUNT 272 a 178 c 174 g 285 t
ORIGIN

```

```

Query Match 25.6%; Score 488.2; DB 10; Length 909;
Best Local Similarity 75.4%; Pred. No. 1.2e-38;
Matches 673; Conservative 0; Mismatches 213; Indels 7; Gaps 5;

```

```

Qy      574 AAAGCACATTTTCATCGTTTGGACAAATTCATCTAAATCAATCCATGCGCCAAAGT 633
Db      2 AAAGCACCTTTTCATCGATTTGACAAATTCATCTAAATCAATCTTGTGGGCAAGC 61
Qy      634 AGGCTCAGAGAAATTTTCTCAACAAAGATATCTTATTCAAGCGCGTTTCTTCTGCTGAG 693
Db      62 AGACTTAGAGATCTTTTAAAGCAGGACATCTTTTCCAGGACGTTTCTGCGAGAA 121
Qy      694 TTGACAAAGCAAGTTTCTCTGACCTTTCTGTCAGCAAAATATCAGATGCGCAGAAATATAGG 753

```

```

Db      122 GTTCAAAAGCAAGTGTCTATCAGATCTTGAACAAGCAAAATATCAGATGGCAGAGTACAGG 181
Qy      754 ATTTCAATCTACGGAAGGAAACAGAGTGAATGGGACCAAACTTGGAAAGTTGGATAGTGAAC 813
Db      182 GTGTCATATATGGAAGGAAACAAAGTGAATGGGACCAAGTTGGCAAGTTGGTTATTATAC 241
Qy      814 AATGAATGCAAGTGGAAATGTTGTCTGGCTGGTTGATGATGATGATGATGATGATGATGATG 873
Db      242 AATGAAATTTATAGTGAGACTACTGTTTGGTTAATCCAGTTTACCAAGATTTGATAATGTG 301
Qy      874 TACAAGGAATGGGTATCGTTACATCAATCCAAATCTTCTTGACAAATTTTGGTCTT 933
Db      302 TACAAGCAATGGGAATTTGTAATATCTTTTTCAGAAATATCTTGGCAATATTTTCAATCCA 361
Qy      934 CTTTTTGAGGTTACTATTATGATCCAGCTTTCACACCCACAGCTCCATGTCTTCTCTGAAGCAG 993
Db      362 CTATTTGAAGTTACAGTTGATCCAAATTTCTCATCTCAACTACACGTTGTTCTTAAATATG 421
Qy      994 GTTGTAGGTTGGACCTGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1053
Db      422 GTGGTTGGCTTTGATTTAGTTGATGATGATGATGATGATGATGATGATGATGATGATG 481
Qy      1054 ATGCCCAACCTGACAGTGGACCAATGTGTTTCAACCTGCAATTTTCATATATTCGCTAC 1113
Db      482 ATGCCAACTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 541
Qy      1114 TACTGCTATGCTAACTTATTCACCTTAAACCAAGCTCGTGTGATGATGATGATGATGATG 1173
Db      542 TACTTTTATGCAAACTTTTATACACTCAACCAAGCTTCTGAGTCAAAAGGATGACAGAA 601
Qy      1174 ATCAAAATTCCTCCACATGCTGGAGAGCTGGAGATGTTGATCACTTGGCAGCGACATTT 1233
Db      602 ATAAACTCCGACCTCACTGTGGGAGGCTGGTGATATGACCATTTAGCTGCTGCTTTC 661
Qy      1234 CTTCTCTGTCAACATATCAGATGGAATTAATCTAAGGAAGTCTCTGTGCTTTCAGTAC 1293
Db      662 CTTCTGTG-CACAATATATCTCACGGGATTAATCTCGGAAATCCCTGTTTTCAGTAC 720
Qy      1294 TTGTACTATCTTGGTCAGATGTTGCTGCGATGTCCTCCCATTTGAGCAACAACCTCCTTATT 1353
Db      721 TTGTATTACCTCGCTCAGATCGGATTTGGCCATGTCCTCTGAGCAATATTTCCCTTTTG 780
Qy      1354 CTTGACTATCATCGCAACCCCTTTTCCAAAGTCTTCTTCCAAAGAGGTCTGAATGTCTCATTA 1413
Db      781 CTGGACTTTTATTGG--ACCCCTTTTCTTCAATCTT-AAACGTGGGCTAAATGTTGCTCTT 837
Qy      1414 TCTAGGATGACCTTTTGCATAATTCACCTGACAAAGAACCAATTTGGTGGGAAGA 1466
Db      838 TCT-CTGATGACCTTTTACAA--TTCTTTGACAAGAGAACCTTTTGGGGAATA 887

```

Search completed: December 15, 2003, 21:05:43
Job time : 3676 secs

GenCore version 5.1.1.6
Copyright. (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 13:20:48 ; Search time 46 Seconds

(without alignments)
1994.434 Million cell updates/sec

Title: US-10-019-633-2

Perfect score: 3100

Sequence: 1 PRVRVAPWEKEVINDPCTPK.....NEMKLIVSYDNEILIPDELIDL 578

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq 19Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3100	100.0	578	22	AA19831948
2	3082	99.4	603	22	AA19831954
3	2683.5	86.6	681	22	AA19831955
4	2644.5	85.3	839	20	AA193456
5	2644.5	85.3	839	23	AA193456
6	2644.5	85.3	859	20	AA193457
7	2342	75.5	595	22	AA19831956
8	2338.5	75.4	492	22	AA19831949
9	1706.5	55.0	345	22	AA19831957

10	1513	52.0	658	22	AB155426	Drosophila melanog
11	1502.5	48.5	747	23	AB170650	Rabbit adenosine m
12	1497	48.3	1813	22	AB170650	Novel human diagn
13	1357.5	43.8	292	23	AA101694	A thaliana AMP dea
14	627	20.2	888	22	AA170751	S cerevisiae apopt
15	404	13.0	82	22	AA193195	Amino acid sequenc
16	344.5	11.1	107	22	AA193195	Amino acid sequenc
17	293	9.5	100	23	AB195952	Human ORF protein
18	191.5	6.2	160	22	AA170501	Human colon cancer
19	184	5.9	65	22	AB150907	Human liver peptid
20	184	5.9	65	22	AB130880	Peptide #3531 enco
21	184	5.9	65	22	AA156848	Human brain expres
22	184	5.9	65	22	AA192332	Human bone marrow
23	184	5.9	65	22	AB138841	Human peptide enco
24	142	4.6	2138	23	AA152295	Human nucleic acid
25	141	4.5	293	22	AB112977	Novel human diagn
26	134	4.3	553	22	AB111988	Drosophila melanog
27	126	4.1	467	22	AB171743	Drosophila melanog
28	120.5	3.9	352	23	AB153593	Lactococcus lactis
29	113.5	3.7	370	22	AB154809	Drosophila melanog
30	111	3.6	1498	22	AB154857	Drosophila melanog
31	111	3.6	1498	23	AA151146	Fruit fly ARMS pro
32	107	3.5	782	11	AA106991	Polypeptide antige
33	106.5	3.4	999	18	AA126682	Bovine lysosomal a
34	106.5	3.4	999	18	AA126684	Bovine lysosomal a
35	106.5	3.4	1032	23	AA121588	Murine Toll-like r
36	105	3.4	464	13	AA129651	AmEPV Spheroidin a
37	105	3.4	464	19	AA141297	AmEPV entomopoxvir
38	105	3.4	464	20	AA130165	Protein encoded by
39	105	3.4	464	23	AB110372	Amino acid encoded
40	105	3.4	464	23	AB177627	AmEPV core protein
41	105	3.4	627	23	AB155152	Lactococcus lactis
42	104.5	3.4	511	19	AA170287	Dendritic cell-der
43	104.5	3.4	667	23	AB192524	Herbicidally activ
44	104.5	3.4	773	22	AB166567	Drosophila melanog
45	104.5	3.4	773	22	AB170352	Drosophila melanog

ALIGNMENTS

RESULT 1

AA19831948

ID AA19831948 standard; Protein; 578 AA.

XX

AC AA19831948;

XX

DT 15-MAY-2001 (first entry)

XX

DE Amino acid sequence of a corn AMP deaminase enzyme.

XX AMP deaminase; adenosine deaminase; adenosine; transgenic plant;

KW inosine.

XX

OS Zea mays.

XX

PN WO200109305-A2.

XX

PD 08-FEB-2001.

XX

PF 28-JUL-2000; 2000MO-US21009.

XX

PR 30-JUL-1999; 99US-0146473.

XX

PA (DUPO) DU PONT DE NEMOURS & CO E I.

PA (PION-) PIONEER HI-BRED INT INC.

XX

PI Caspar T, Falco SC, Sakai H, Weng Z, Hu X;

XX

DR WPI; 2001-159866/16.

DR N-PSDB; AAF25488.

XX

PT New polynucleotides encoding AMP deaminase or adenosine deaminase used

PT e.g. for designing or identifying herbicides that inhibit the enzyme
PT activities, and as probes for genetic or physical mapping -
XX
PS Claim 10; Page 43-44; 72pp; English.

XX The present sequence represents an AMP deaminase. The specification also
CC describes adenosine deaminase. These enzymes convert adenosine to
CC inosine. Mutations in these genes cause disruptions in then salvage and
CC catabolism of adenosine and AMP. In humans, this may lead to death of
CC white blood cells, which causes severe immunodeficiencies. The AMP
CC deaminase and adenosine deaminase may be used to prepare antibodies
CC to these proteins, and to design or identify herbicides that inhibit
CC their enzyme activities. The polynucleotides are used as probes for
CC genetically and physically mapping genes that they compose, and as
CC markers for traits linked to those genes, where such information may be
CC used in plant breeding to develop lines with desired phenotypes. The
CC nucleic acid fragments may be used to isolate cDNAs and genes encoding
CC homologous proteins from the same or other plant species, and in
CC polymerase chain reaction (PCR) protocols to amplify longer nucleic
CC acid fragments encoding homologous genes from DNA or RNA. These may
CC also be used to create transgenic plants in which the polypeptides are
CC overexpressed or suppressed, and as probes in direct fluorescent in
CC situ hybridisation (FISH).

XX Sequence 578 AA;

Query Match 100.0%; Score 3100; DB 22; Length 578;
Best Local Similarity 100.0%; Pred. No. 1.3e-299;
Matches 578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRVRAPWEKEVINDCTPKPNPFTYVPEPKSEHVFTQDGVHIVYADKCTESIYPV 60

Db 1 PRVRAPWEKEVINDCTPKPNPFTYVPEPKSEHVFTQDGVHIVYADKCTESIYPV 60

QY 61 ADATFTFDLHYLTRVTAAGNTRTVCHNRLNLEHFKFHLMLNADREFLAQTAHPRDF 120

Db 61 ADATFTFDLHYLTRVTAAGNTRTVCHNRLNLEHFKFHLMLNADREFLAQTAHPRDF 120

QY 121 YNVRKVDTHVHSACMNQKHLRFKSKURKEPDEVVIFRDGTMTLKEVFESLDTGYD 180

Db 121 YNVRKVDTHVHSACMNQKHLRFKSKURKEPDEVVIFRDGTMTLKEVFESLDTGYD 180

QY 181 LVNVDLDDVHADKSTFHRFDKFNLYKPCQOSRLREIFLQDNLIQGRFLAELTKQVFSDL 240

Db 181 LVNVDLDDVHADKSTFHRFDKFNLYKPCQOSRLREIFLQDNLIQGRFLAELTKQVFSDL 240

QY 241 SASKYQMAEYRISYIGRKQSEWDQLASWIVNNELHSGNVVWLQVLPRLYNYKEMGIVTS 300

Db 241 SASKYQMAEYRISYIGRKQSEWDQLASWIVNNELHSGNVVWLQVLPRLYNYKEMGIVTS 300

QY 301 FQNLNLI FVPLFEVTDIDPASHPOLHVLKQVVGDLVDDESKPERPPTKHMTPEQWTN 360

Db 301 FQNLNLI FVPLFEVTDIDPASHPOLHVLKQVVGDLVDDESKPERPPTKHMTPEQWTN 360

QY 361 VFNPAFSYAYCYANLFTLNKLRKSGMTTIKFRPHAGEAGVDVHLAATFLCHNISHG 420

Db 361 VFNPAFSYAYCYANLFTLNKLRKSGMTTIKFRPHAGEAGVDVHLAATFLCHNISHG 420

QY 421 INLRKSPVLYLYLQOIGLAMSPLNNSLFLDYHRNPFTTFFQGLNVLSTDDPLQIH 480

Db 421 INLRKSPVLYLYLQOIGLAMSPLNNSLFLDYHRNPFTTFFQGLNVLSTDDPLQIH 480

QY 481 LTKEPLVEEYSIAASLWKLSSCDLCEIARNSVYQSGFSHALKAHWIGKVKRGPAGNDI 540

Db 481 LTKEPLVEEYSIAASLWKLSSCDLCEIARNSVYQSGFSHALKAHWIGKVKRGPAGNDI 540

QY 541 HRTNVPHIRVQFREMIVRNEMKLVYSDNEILLIPDELDL 578

Db 541 HRTNVPHIRVQFREMIVRNEMKLVYSDNEILLIPDELDL 578

RESULT 2
AAB31954

ID AAB31954 standard; Protein; 603 AA.
XX AAB31954;
XX
DT 15-MAY-2001 (first entry)
DE Amino acid sequence of a soybean AMP deaminase enzyme.
XX AMP deaminase; adenosine deaminase; adenosine; transgenic plant;
KW inosine.
XX Glycine max.
XX Key Location/Qualifiers
FT Misc-difference 15
FT /note= "unknown residue encoded by CAG"
XX
PN WO200109305-A2.
XX 08-FEB-2001.
XX 28-JUL-2000; 2000WO-US21009.
XX 30-JUL-1999; 99US-0146473.
XX (DUPO) DU PONT DE NEMOURS & CO E I.
PA (PION-) PIONEER HI-BRED INT INC.
XX Caspar T, Falco SC, Sakai H, Weng Z, Hu X;
PI
DR WPI; 2001-159866/16.
XX N-PSDB; AAF25494.

PT New polynucleotides encoding AMP deaminase or adenosine deaminase used
PT e.g. for designing or identifying herbicides that inhibit the enzyme
PT activities, and as probes for genetic or physical mapping -
XX
PS Claim 10; Page 57-59; 72pp; English.

XX The present sequence represents an AMP deaminase. The specification also
CC describes adenosine deaminase. These enzymes convert adenosine to
CC inosine. Mutations in these genes cause disruptions in then salvage and
CC catabolism of adenosine and AMP. In humans, this may lead to death of
CC white blood cells, which causes severe immunodeficiencies. The AMP
CC deaminase and adenosine deaminase may be used to prepare antibodies
CC to these proteins, and to design or identify herbicides that inhibit
CC their enzyme activities. The polynucleotides are used as probes for
CC genetically and physically mapping genes that they compose, and as
CC markers for traits linked to those genes, where such information may be
CC used in plant breeding to develop lines with desired phenotypes. The
CC nucleic acid fragments may be used to isolate cDNAs and genes encoding
CC homologous proteins from the same or other plant species, and in
CC polymerase chain reaction (PCR) protocols to amplify longer nucleic
CC acid fragments encoding homologous genes from DNA or RNA. These may
CC also be used to create transgenic plants in which the polypeptides are
CC overexpressed or suppressed, and as probes in direct fluorescent in
CC situ hybridisation (FISH).

XX Sequence 603 AA;

Query Match 99.4%; Score 3082; DB 22; Length 603;
Best Local Similarity 99.7%; Pred. No. 8.8e-298;
Matches 575; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 RVRVAPWEKEVINDCTPKPNPFTYVPEPKSEHVFTQDGVHIVYADKCTESIYPVA 61

Db 27 REEVAPWEKEVINDCTPKPNPFTYVPEPKSEHVFTQDGVHIVYADKCTESIYPVA 86

QY 62 DATTEFTDLHYLTRVTAAGNTRTVCHNRLNLEHFKFHLMLNADREFLAQTAHPRDF 121

Db 87 DATTEFTDLHYLTRVTAAGNTRTVCHNRLNLEHFKFHLMLNADREFLAQTAHPRDF 146

QY 122 NVVRKVDTHVHSACMNQKHLRFKSKURKEPDEVVIFRDGTMTLKEVFESLDTGYD 181

Db 147 NVRKVDTHVHHSACNMQRHLLRFKSLRKEPDEVIFRDGYMTLKEVFSLDLTGYDL 206
Qy 182 NVDLLDVHADKSTFHRDFKFNLYKPCGQSRLEIFLKQDNLIQGRFLAELTKQVFSDL 241
Db 207 NVDLLDVHADKSTFHRDFKFNLYKPCGQSRLEIFLKQDNLIQGRFLAELTKQVFSDL 266
Qy 242 ASKYQMAEYRISYGRKQSEWDLASWVNNELHSGNVVWLQVLPRLYNYKEMGIVTSF 301
Db 267 ASKYQMAEYRISYGRKQSEWDLASWVNNELHSGNVVWLQVLPRLYNYKEMGIVTSF 326
Qy 302 QNLLDNIFVLPFEVITDPASHPOLHVLKQVVGDLVDDESXPERPRTKHMPTPQWNTV 361
Db 327 QNLLDNIFVLPFEVITDPASHPOLHVLKQVVGDLVDDESXPERPRTKHMPTPQWNTV 386
Qy 362 FNPAFSYAYCYANLFTLNKLRKSGMTTIKFRPHAGEAGVDVHLAATFLLCHNISHGI 421
Db 387 FNPAFSYAYCYANLFTLNKLRKSGMTTIKFRPHAGEAGVDVHLAATFLLCHNISHGI 446
Qy 422 NLRKSPVLQYLYLGOIGLAMSPLSNNSLFLDYHRNPFTTQGRGLNVSLSLTDPLQIHL 481
Db 447 NLRKSPVLQYLYLGOIGLAMSPLSNNSLFLDYHRNPFTTQGRGLNVSLSLTDPLQIHL 506
Qy 482 TKEPLVEEYSIAASLWKLSSCDLCEIARNVYQSGFSHALKAHWIGKNYFKRGPAGNDIH 541
Db 507 TKEPLVEEYSIAASLWKLSSCDLCEIARNVYQSGFSHALKAHWIGKNYFKRGPAGNDIH 566
Qy 542 RTNVPHIRVQFREMIVRNEMKLVSDNEILIPDEL 578
Db 567 RTNVPHIRVQFREMIVRNEMKLVSDNEILIPDEL 603

RESULT 3

AAAB31955
ID AAB31955 standard; Protein; 681 AA.
XX
AC AAB31955;
XX
DT 15-MAY-2001 (first entry)
XX
DE Amino acid sequence of a wheat AMP deaminase enzyme.
XX
KW AMP deaminase; adenosine deaminase; adenosine; transgenic plant;
KW inosine.
XX
OS Triticum aestivum.
XX
PN WO200109305-A2.
XX
PD 08-FEB-2001.
XX
PF 28-JUL-2000; 2000WO-US21009.
XX
PR 30-JUL-1999; 99US-0146473.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX (PION-) PIONEER HI-BRED INT INC.
XX
PI Caspar T, Falco SC, Sakai H, Weng Z, Hu X;
XX
XX WPI; 2001-159866/16.
DR N-PSDB; AAF25495.
XX
PT New polynucleotides encoding AMP deaminase or adenosine deaminase used
PT e.g. for designing or identifying herbicides that inhibit the enzyme
PT activities, and as probes for genetic or physical mapping
XX
PS Claim 10; Page 60-62; 72pp; English.
XX
XX The present sequence represents an AMP deaminase. The specification also
CC describes adenosine deaminase. These enzymes convert adenosine to
CC inosine. Mutations in these genes cause disruptions in then salvage and
CC catabolism of adenosine and AMP. In humans, this may lead to death of

CC white blood cells, which causes severe immunodeficiencies. The AMP
CC deaminase and adenosine deaminase may be used to prepare antibodies
CC to these proteins, and to design or identify herbicides that inhibit
CC their enzyme activities. The polynucleotides are used as probes for
CC genetically and physically mapping genes that they compose, and as
CC markers for traits linked to those genes, where such information may be
CC used in plant breeding to develop lines with desired phenotypes. The
CC nucleic acid fragments may be used to isolate cDNAs and genes encoding
CC homologous proteins from the same or other plant species, and in
CC polymerase chain reaction (PCR) protocols to amplify longer nucleic
CC acid fragments encoding homologous genes from DNA or RNA. These may
CC also be used to create transgenic plants in which the polypeptides are
CC overexpressed or suppressed, and as probes in direct fluorescent in
CC situ hybridisation (FISH).
XX
SQ Sequence 681 AA;
Query Match 86.6%; Score 2683.5; DB 22; Length 681;
Best Local Similarity 85.1%; Pred. No. 5.6e-259;
Matches 490; Conservative 42; Mismatches 43; Indels 1; Gaps 1;
Qy 2 RVRVAPWEKVINDPCTPKPNPNPTVYVPEKSEHVFOVDGVHVVADKDCSTIYPA 61
Db 106 REEVAPWEKEIITDPTPKPNPNPTVYVPEKSEHVFOVDGVHVVADKDCSTIYPA 165
Qy 62 DATTFPTDLHYLRVTAAGNTRTVCHNRLNLEHKEFHLMLNADREFLAQKTAPHRDFY 121
Db 166 DATTFPTDMVILRVLAAGDIRTVCYKRLNLEHKEFHLMLNADREFLAQKTAPHRDFY 225
Qy 122 NVRKVDTHVHHSACNMQRHLLRFKSLRKEPDEVIFRDGYMTLKEVFSLDLTGYDL 181
Db 226 NVRKVDTHVHHSACNMQRHLLRFKSLRKEPDEVIFRDGYMTLKEVFSLDLTGYDL 285
Qy 182 NVDLLDVHADKSTFHRDFKFNLYKPCGQSRLEIFLKQDNLIQGRFLAELTKQVFSDL 241
Db 286 NVDLLDVHADKSTFHRDFKFNLYKPCGQSRLEIFLKQDNLIQGRFLAELTKQVFSDL 345
Qy 242 ASKYQMAEYRISYGRKQSEWDLASWVNNELHSGNVVWLQVLPRLYNYKEMGIVTSF 301
Db 346 ASKYQMAEYRISYGRKQSEWDLASWVNNELHSGNVVWLQVLPRLYNYKEMGIVTSF 405
Qy 302 QNLLDNIFVLPFEVITDPASHPOLHVLKQVVGDLVDDESXPERPRTKHMPTPQWNTV 361
Db 406 QNLLDNIFVLPFEVITDPASHPOLHVLKQVVGDLVDDESXPERPRTKHMPTPQWNTV 465
Qy 362 FNPAFSYAYCYANLFTLNKLRKSGMTTIKFRPHAGEAGVDVHLAATFLLCHNISHGI 421
Db 466 FNPAYAYVYCYANLFTLNKLRKSGMTTIKFRPHAGEAGVDVHLAATFLLCHNISHGI 525
Qy 422 NLRKSPVLQYLYLGOIGLAMSPLSNNSLFLDYHRNPFTTQGRGLNVSLSLTDPLQIHL 481
Db 526 NLRKSPVLQYLYLGOIGLAMSPLSNNSLFLDYHRNPFTTQGRGLNVSLSLTDPLQIHL 585
Qy 482 TKEPLVEEYSIAASLWKLSSCDLCEIARNVYQSGFSHALKAHWIGKNYFKRGPAGNDIH 541
Db 586 TKEPLVEEYSIAASLWKLSSCDLCEIARNVYQSGFSHALKAHWIGKNYFKRGPAGNDIH 645
Qy 542 RTNVPHIRVQFREMIVRNEMKLVSDNEILIPDEL 577
Db 646 QTNVPHIRIEPRHTIWKEMELIHLRN-VDIPEID 680
RESULT 4
AAAY33456
ID AAY33456 standard; Protein; 839 AA.
XX
AC AAY33456;
XX
DT 13-DEC-1999 (first entry)
XX
DE A. thaliana AMP-deaminase protein #1.
XX
KW AMP-deaminase; adenosine triphosphate aminohydrolase; plant; herbicide;

Db 387 KVDTHVHSCAMQKELLRFIISKLRKEPDEVIFRDGTYLTREVFESDLDTGYDLNVD 446
 QY 185 LLDVHADKSTFHRDFKFNLYNPKCGOSRLREIFLKQDNLIQGRFLAELTKQVFSLSASK 244
 Db 447 LLDVHADKSTFHRDFKFNLYNPKCGOSRLREIFLKQDNLIQGRFLGELTKQVFSLSASK 506
 QY 245 YQMAEYRISIGRKQSEWDQLASWIVNNELHSGNVLMVQIPLRYNYKEMGIVTSFQNL 304
 Db 507 YQMAEYRISIGRKQSEWDQLASWIVNNELHSGNVLMVQIPLRYNYKEMGIVTSFQNI 566
 QY 305 LDNIFFPLFEVITIDPASHQVFLKQVVGDLVDDESKPERPPTKHMPTPEQWTVNPNP 364
 Db 567 LDNIFFPLFEVITIDPASHQVFLKQVVGDLVDDESKPERPPTKHMPTPEQWTVNPNP 626
 QY 365 AFSYVYVYCYANLFTLNKLRSGKMTTIFRPHAGEAGDVHDLAATFLCHNISHGINLR 424
 Db 627 AFSYVYVYCYANLFTLNKLRSGKMTTITLRPHSGEAGDIDHLAATFLCHNISHGINLR 686
 QY 425 KSPVLYLYLQIGLAMSPLSNNSLFLDYHRNPPTFFQRLGNVSLSTDDPLQIHLTKE 484
 Db 687 KSPVLYLYLQIGLAMSPLSNNSLFLDYHRNPPTFFQRLGNVSLSTDDPLQIHLTKE 746
 QY 485 PLVEEYSIAASLWKLSSCDLCEIARNVYQSGFSGFHALKAHWIGKGYFKGPGAGNDIHRTN 544
 Db 747 PLVEEYSIAASVWKLSSCDLCEIARNVYQSGFSGFHALKSHWIGKDYKRGPGDNDIHKTN 806
 QY 545 VPHIRVOFREMIRNEMKLVYSDNEILIPDEL 576
 Db 807 VPHIRVEFRDTIWKEMQQVYL-GKAVISDEV 837

RESULT 6

AAY33457
 ID AAY33457 standard; Protein; 859 AA.

AC AAY33457;

DT 13-DEC-1999 (first entry)

DE A. thaliana AMP-deaminase protein #2.

KW AMP-deaminase; adenosine triphosphate aminohydrolase; plant; herbicide;
 KW inhibitor; resistance.

OS Arabidopsis thaliana.

PN W09950400-A1.

XX 07-OCT-1999.

PF 25-MAR-1999; 99WO-EP02016.

PR 01-APR-1998; 98DE-1014512.

PA (BADI) BASF AG.

XX Lerchl J, Reindl A;

XX WPI; 1999-580759/49.

DR N-PSDB; AAZ23395.

XX DNA encoding AMP deaminase, assay systems for identifying inhibitors

PT and transgenic plants -

XX Disclosure; Fig 2; 45pp; German.

CC This invention describes a novel Arabidopsis thaliana AMP-deaminase
 CC (adenosine triphosphate aminohydrolase, EC 3.5.4.6) which has herbicidal
 CC activity. The AMP deaminase DNA sequence, can be introduced into pro- or
 CC eukaryotic cells, with the relevant control elements to control
 CC transcription and translation in the cell. An expression cassette
 CC derived from the products of the invention is used to transform plants.

CC It is useful for producing a test system to identify inhibitors of AMP
 CC deaminase. Plants containing the expression cassette are useful for the
 CC production of the AMP deaminase. The plants have an increased resistance
 CC against inhibitors of the AMP-deaminase through strengthened expression
 CC of the AMP deaminase DNA sequence. The expression cassette is useful for
 CC producing plants with an increased content of AMP. This sequence
 CC represents the A. thaliana AMP-deaminase described in the invention.

SQ Sequence 859 AA;

Query Match 85.3%; Score 2644.5; DB 20; Length 859;

Best Local Similarity 83.9%; Pred. No. 6.2e-254; Indels 1; Gaps 1;

Matches 480; Conservative 46; Mismatches 45;

QY 5 VAPKEKEVINDPCTPKPNPNFTYVPEPKSEHVFTQVDGVIHVADKCTESIYVADAT 64

Db 287 VAPKEKEVISDPSIPKNTPEFAHYPOGKSDHCEMQDGVVHVFAKDAKEDLFFVADAT 346

QY 65 TFFTDLHVILRVTAAGNRTVCHNRNLNLEHKKFHLMLNADREFLAOKTAPHRDFYNVR 124

Db 347 AFTDHLHVLVIAAGNIRTLCHRRLLVLEQKFNHLMLNADKEFLAOKSAPHRDFYNVR 406

QY 125 KVDTHVHSCAMQKELLRFIISKLRKEPDEVIFRDGTYLTREVFESDLDTGYDLNVD 184

Db 407 KVDTHVHSCAMQKELLRFIISKLRKEPDEVIFRDGTYLTREVFESDLDTGYDLNVD 466

QY 185 LLDVHADKSTFHRDFKFNLYNPKCGOSRLREIFLKQDNLIQGRFLAELTKQVFSLSASK 244

Db 467 LLDVHADKSTFHRDFKFNLYNPKCGOSRLREIFLKQDNLIQGRFLGELTKQVFSLSASK 526

QY 245 YQMAEYRISIGRKQSEWDQLASWIVNNELHSGNVLMVQIPLRYNYKEMGIVTSFQNL 304

Db 527 YQMAEYRISIGRKQSEWDQLASWIVNNELHSGNVLMVQIPLRYNYKEMGIVTSFQNI 586

QY 305 LDNIFFPLFEVITIDPASHQVFLKQVVGDLVDDESKPERPPTKHMPTPEQWTVNPNP 364

Db 587 LDNIFFPLFEVITIDPASHQVFLKQVVGDLVDDESKPERPPTKHMPTPEQWTVNPNP 646

QY 365 AFSYVYVYCYANLFTLNKLRSGKMTTIFRPHAGEAGDVHDLAATFLCHNISHGINLR 424

Db 647 AFSYVYVYCYANLFTLNKLRSGKMTTITLRPHSGEAGDIDHLAATFLCHNISHGINLR 706

QY 425 KSPVLYLYLQIGLAMSPLSNNSLFLDYHRNPPTFFQRLGNVSLSTDDPLQIHLTKE 484

Db 707 KSPVLYLYLQIGLAMSPLSNNSLFLDYHRNPPTFFQRLGNVSLSTDDPLQIHLTKE 766

QY 485 PLVEEYSIAASLWKLSSCDLCEIARNVYQSGFSGFHALKAHWIGKGYFKGPGAGNDIHRTN 544

Db 767 PLVEEYSIAASVWKLSSCDLCEIARNVYQSGFSGFHALKSHWIGKDYKRGPGDNDIHKTN 826

QY 545 VPHIRVOFREMIRNEMKLVYSDNEILIPDEL 576

Db 827 VPHIRVEFRDTIWKEMQQVYL-GKAVISDEV 857

RESULT 7

AAB31956

ID AAB31956 standard; Protein; 595 AA.

XX AAB31956;

AC AAB31956;

XX 15-MAY-2001 (first entry)

DT Amino acid sequence of a corn adenosine deaminase enzyme.

DE AMP deaminase; adenosine deaminase; adenosine; transgenic plant;

KW inosine.

XX Zea mays.

OS WO200109305-A2.

XX 08-FEB-2001.

```
XX PF 28-JUL-2000; 2000WO-US21009.
XX PR 30-JUL-1999; 99US-0146473.
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX PA (PION-) PIONEER HI-BRED INT INC.
XX PI Caspar T, Falco SC, Sakai H, Weng Z, Hu X;
XX DR WPI; 2001-159866/16.
XX DR N-PSDB; AAF25496.
XX PT New polynucleotides encoding AMP deaminase or adenosine deaminase used
XX PT e.g. for designing or identifying herbicides that inhibit the enzyme
XX PT activities, and as probes for genetic or physical mapping
XX PS Claim 10; Page 63-65; 72pp; English.
XX CC The present sequence represents adenosine deaminase. The specification
XX CC also describes adenosine deaminase. These enzymes convert adenosine to
XX CC inosine. Mutations in these genes cause disruptions in then salvage and
XX CC catabolism of adenosine and AMP. In humans, this may lead to death of
XX CC white blood cells, which causes severe immunodeficiencies. The AMP
XX CC deaminase and adenosine deaminase may be used to prepare antibodies
XX CC to these proteins, and to design or identify herbicides that inhibit
XX CC their enzyme activities. The polynucleotides are used as probes for
XX CC genetically and physically mapping genes that they compose, and as
XX CC markers for traits linked to those genes, where such information may be
XX CC used in plant breeding to develop lines with desired phenotypes. The
XX CC nucleic acid fragments may be used to isolate cDNAs and genes encoding
XX CC homologous proteins from the same or other plant species, and in
XX CC polymerase chain reaction (PCR) protocols to amplify longer nucleic
XX CC acid fragments encoding homologous genes from DNA or RNA. These may
XX CC also be used to create transgenic plants in which the polypeptides are
XX CC overexpressed or suppressed, and as probes in direct fluorescent in
XX CC situ hybridisation (FISH).
XX SQ Sequence 595 AA;
Query Match 75.5%; Score 2342; DB 22; Length 595;
Best Local Similarity 87.7%; Pred. No. 5e-224;
Matches 429; Conservative 32; Mismatches 28; Indels 0; Gaps 0;
QY 2 RVAVAPWEKVINDPCTKPNPNFTYVPEPKSEHVFTQVDGVIHVYADKDCETSIYPA 61
DB 107 REAVAPWDKEVYSDPSTPKPNPDFFLYIPEGNSDHYFEMQDGVIRYVPRDAKEELFPA 166
QY 62 DATFTFFDLHVLIRVTAAGNTRTVCHNRLNLLHKKFKFLMLNADREFLAQKTAHRDFY 121
DB 167 DATFTFFDLHLLRVIAAGNIRILCHHRLNLLSQKNFLNMLNADREFLAQKSAHRDFY 226
QY 122 NVRKVDTHVHHSACMNQKHLRFIKSKRKEPDEVVIFRDGTMTLKEVPESLDLTGYDL 181
DB 227 NVRKVDTHVHHSACMNQKHLRFIKSKRKEPDEVVIFRDGTMTLKEVPESLDLTGYDL 286
QY 182 NVDLLDVHADKSTFHRFDKFNLYNCCGQSRRLREIFLKQDNLTQGRFLAELTKQVSDLS 241
DB 287 NVDLLDVHADKSTFHRFDKFNLYNCCGQSRRLREIFLKQDNLTQGRFLAELTKQVSDLS 346
QY 242 ASKYQMAEYRISYIGRKQSEWDLASWVNNELHSGNVWLVQIPRLXNYVKEMGIVTSF 301
DB 347 ASKYQMAEYRISYIGRKQSEWDLASWVNNELHSGNVWLVQIPRLXNYVKEMGIVTSF 406
QY 302 QNLLDNIFVLPFVETIDPASHPOLHVFLKQVGLDLVDDESKPERPRTKHMPTPEQWTVN 361
DB 407 QNMLDNIFLPFVETVNPDPSPHQLHVFLKQVGLDLVDDESKPERPRTKHMPTPEQWTVN 466
QY 362 FNPAFSYAYCYANLFTLNKLRSGKMTIKRPHAGEAGDVLDHAAATPELLCHNISHGI 421
DB 467 FNPAFSYAYCYANLFTLNKLRSGKMTIKRPHSGEAGDLDHAAATPELLTAHNAHGI 526
QY 422 NLRKSPVLQYLYLQIGLAMSPLSNNSFLDYHRNPFPPFGLNLSLSTDDPLQIHL 481
DB 527 NLRKSPVLQYLYLQIGLAMSPLSNNSFLDYHRNPFPPFGLNLSLSTDDPLQIHL 586
QY 482 TKEPLVEEY 490
DB 587 TKEPLVEEY 595
RESULT 8
AAB31949
ID AAB31949 standard; Protein; 492 AA.
AC AAB31949;
XX 15-MAY-2001 (first entry)
XX Amino acid sequence of a rice AMP deaminase enzyme.
XX AMP deaminase; adenosine deaminase; adenosine; transgenic plant;
XX inosine.
XX Oryza sativa.
XX WO200109305-A2.
XX 08-FEB-2001.
XX 28-JUL-2000; 2000WO-US21009.
XX 30-JUL-1999; 99US-0146473.
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX (PION-) PIONEER HI-BRED INT INC.
XX Caspar T, Falco SC, Sakai H, Weng Z, Hu X;
XX WPI; 2001-159866/16.
XX DR N-PSDB; AAF25489.
XX New polynucleotides encoding AMP deaminase or adenosine deaminase used
XX e.g. for designing or identifying herbicides that inhibit the enzyme
XX activities, and as probes for genetic or physical mapping
XX Claim 10; Page 45-47; 72pp; English.
XX The present sequence represents an AMP deaminase. The specification also
XX describes adenosine deaminase. These enzymes convert adenosine to
XX inosine. Mutations in these genes cause disruptions in then salvage and
XX catabolism of adenosine and AMP. In humans, this may lead to death of
XX white blood cells, which causes severe immunodeficiencies. The AMP
XX deaminase and adenosine deaminase may be used to prepare antibodies
XX to these proteins, and to design or identify herbicides that inhibit
XX their enzyme activities. The polynucleotides are used as probes for
XX genetically and physically mapping genes that they compose, and as
XX markers for traits linked to those genes, where such information may be
XX used in plant breeding to develop lines with desired phenotypes. The
XX nucleic acid fragments may be used to isolate cDNAs and genes encoding
XX homologous proteins from the same or other plant species, and in
XX polymerase chain reaction (PCR) protocols to amplify longer nucleic
XX acid fragments encoding homologous genes from DNA or RNA. These may
XX also be used to create transgenic plants in which the polypeptides are
XX overexpressed or suppressed, and as probes in direct fluorescent in
XX situ hybridisation (FISH).
XX SQ Sequence 492 AA;
Query Match 75.4%; Score 2338.5; DB 22; Length 492;
Best Local Similarity 87.3%; Pred. No. 8.2e-224;
Matches 427; Conservative 35; Mismatches 26; Indels 1; Gaps 1;
QY 89 RLNLLEHKFKFLMLNADREFLAQKTAHRDFYVNRKVDTHVHHSACMNQKHLRFIKSK 148
DB 4 RLNLLEKFNLMVNNADRELLAQKAAHRDFYVNRKVDTHVHHSACMNQKHLRFIKSK 63
```


QY 149 LRKEPDEVIFRDGTYMTLKEVPESLDLTGYDLNVLDLVDHAKSTFHRFDKFNLYKNPC 208
 Db 64 LRKEPDEVIFRDGTYMTLKEVPESLDLTGYDLNVLDLVDHAKSTFHRFDKFNLYKNPC 123
 QY 209 GOSRLREIFLKQDNLTQGRFLAELTKQVSDLSASKYQMAEYRISYGRKQSEWDOLASW 268
 Db 124 GOSRLREIFLKQDNLTQGRFLAELTKQVSDLSASKYQMAEYRISYGRKQSEWDOLASW 183
 QY 269 IYNNELHSGNVVLMQIPRLYLVYKEMGIVTSFQNLNDNIFVPLFEVTTIDPASHQPHVF 328
 Db 184 IYNNELHSGNVVLMQIPRIYVYREMGITNSFQNLNDNIFLPLFEVTVDPASHQPHVF 243
 QY 329 LKQVGLDLVDDESKDPERPTKMTPTPEQWTVNFNPAFSYAYCYANLFTLNKLRSGK 388
 Db 244 LKQVGLDLVDDESKDPERPTKMTPTPEQWTVNFNPAFSYAYCYANLFTLNKLRSGK 303
 QY 389 MTTIKFRPHAGAGVDVHLAATFLCHNLSHGINLRKSPVLOVLYLGOIGLAMSPLSN 448
 Db 304 MTTIKLRPHCGAGDIDHLLAAFLTSHNIAHGNLKKSPVLOVLYLGOIGLAMSPLSN 363
 QY 449 SLFLDYHRNPFTFFQRLGNLSLSTDDPLQIHLTKPLVEEYSIAASLWKLSSCDLCEIA 508
 Db 364 SMFIDYHRNPFTFFQRLGNLSLSTDDPLQIHLTKPLVEEYSIAASLWKLSSCDLCEIA 423
 QY 509 RNSVQSGFSLKAKHAWIGKNYKRGPGAGNDIHRNTNVPVPHIRVOFREMIMRNEMKLVYSDN 568
 Db 424 RNSVQSGFSLKAKHAWIGKNYKRGPGAGNDIHRNTNVPVPHIRVOFREMIMRNEMKLVYSDN 483
 QY 569 ELLIPDELD 577
 Db 484 -VDIPEED 491

RESULT 9

AAB31957
 ID AAB31957 standard; Protein; 345 AA.

XX AC AAB31957;

DT 15-MAY-2001 (first entry)

XX Amino acid sequence of a soybean adenosine deaminase enzyme.

XX AMP deaminase; adenosine deaminase; adenosine; transgenic plant;
 KW inosine.

XX Glycine max.

XX Key Location/Qualifiers

FT Misc-difference 4 /note= "Phe encoded by TT"

XX WO200109305-A2.

XX PD 08-FEB-2001.

XX PF 28-JUL-2000; :2000WO-US21009.

XX PR 30-JUL-1999; 99US-0146473.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

PA (PION-) PIONEER HI-BRED INT INC.

XX Caspar T, Falco SC, Sakai H, Weng Z, Hu X;

XX WPI; 2001-159866/16.

DR N-PSDB; AAF25497.

XX New polynucleotides encoding AMP deaminase or adenosine deaminase used
 PT e.g. for designing or identifying herbicides that inhibit the enzyme
 PT activities, and as probes for genetic or physical mapping -

XX

PS Claim 10; Page 66-67; 72pp; English.

XX The present sequence represents adenosine deaminase. The specification
 CC also describes adenosine deaminase. These enzymes convert adenosine to
 CC inosine. Mutations in these genes cause disruptions in then salvage and
 CC catabolism of adenosine and AMP. In humans, this may lead to death of
 CC white blood cells, which causes severe immunodeficiencies. The AMP
 CC deaminase and adenosine deaminase may be used to prepare antibodies
 CC to these proteins, and to design or identify herbicides that inhibit
 CC their enzyme activities. The polynucleotides are used as probes for
 CC genetically and physically mapping genes that they compose, and as
 CC markers for traits linked to those genes, where such information may be
 CC used in plant breeding to develop lines with desired phenotypes. The
 CC nucleic acid fragments may be used to isolate cDNAs and genes encoding
 CC homologous proteins from the same or other plant species, and in
 CC polymerase chain reaction (PCR) protocols to amplify longer nucleic
 CC acid fragments encoding homologous genes from DNA or RNA. These may
 CC also be used to create transgenic plants in which the polypeptides are
 CC overexpressed or suppressed, and as probes in direct fluorescent in
 CC situ hybridisation (FISH).

XX SQ Sequence 345 AA;

Query Match 55.0%; Score 1706.5; DB 22; Length 345;

Best Local Similarity 91.8%; Pred. NO. 4.7e-161;

Matches 313; Conservative 22; Mismatches 5; Indels 1; Gaps 1;

QY 237 FSDLSASKYQMAEYRISYGRKQSEWDOLASWVNNELHSGNVVLMQIPRLYLVYKEMG 296
 Db 5 FSDLSASKYQMAEYRISYGRKQSEWDOLASWVNNELHSGNVVLMQIPRLYLVYKEMG 64

QY 237 IVTSFQNLNDNIFVPLFEVTTIDPASHQPHVFQVGLDLVDDESKDPERPTKMTPTPE 356
 Db 65 IVTSFQNLNDNIFVPLFEVTTIDPASHQPHVFQVGLDLVDDESKDPERPTKMTPTPE 124

QY 357 QWTVNFNPAFSYAYCYANLFTLNKLRSGKMTIKFRPHAGAGVDVHLAATFLCHN 416
 Db 125 QWTVNFNPAFSYAYCYANLFTLNKLRSGKMTIKFRPHAGAGVDVHLAATFLCHN 184

QY 417 ISHGINLRKSPVLOVLYLGOIGLAMSPLSNLSFLDYHRNPFTFFQRLGNLSLSTDDP 476
 Db 195 ISHGINLRKSPVLOVLYLGOIGLAMSPLSNLSFLDYHRNPFTFFQRLGNLSLSTDDP 244

QY 477 LQIHLTKPLVEEYSIAASLWKLSSCDLCEIARNVYQSGFSLKAKHAWIGKNYKRGPA 536
 Db 245 LQIHLTKPLVEEYSIAASLWKLSSCDLCEIARNVYQSGFSLKAKHAWIGKNYKRGPS 304

QY 537 GNDIHRNTNVPVPHIRVOFREMIMRNEMKLVYSDNEILLIPDELD 577

Db 305 GNDIHRNTNVPVPHIRVOFREMIMRNEMKLVYSDNEILLIPDELD 344

RESULT 10

ABB65426

ID ABB65426 standard; Protein; 658 AA.

XX AC ABB65426;

XX DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 23070.

XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US09231.

XX

PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
XX PI
XX N-PSDB; ABL09529.
XX WPI; 2001-656860/75.
XX N-PSDB; ABL09529.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX Disclosure; SEQ ID NO 23070; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABBS7737-ABBS72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 658 AA;
SQ
Query Match 52.0%; Score 1613; DB 22; Length 658;
Best Local Similarity 53.6%; Pred. No. 2.7e-151;
Matches 302; Conservative 93; Mismatches 152; Indels 16; Gaps 5;
QY 22 NNPFTYVPEPKS-----EHVFQTVGV-----IHHVADKCTESI-YPVADATFTT 68
DB 87 NPITYGHKECKAIFWSSLVQWYQLMHGLGWFGLNLNATLDESSEIKIYEYPMDSQFVN 146
QY 69 DLHYILRVTAAGNTRVTCNRLNLEHKKFKHMLNADREFLAQKTAHPRDFVNVKVT 128
DB 147 DMQVCMNMTADGPKSFCYRRLCYLSKSKYQMHVLLNLELAQAQVPHRDFVNTKVT 206
QY 129 HVHSSACMOKHLLRIKSKRKEPDEVIFRDTYMTLKEVPESLDLGYDLNVLDDV 188
DB 207 HIHAACMOKHLLRIKSKRKEPDEVIFRDTYMTLKEVPESLDLGYDLNVLDDV 266
QY 189 HADKSTFHRPDKNLYNCPGQRLREIFLQKDLQGRFLATKQVSDLSASKYQWA 248
DB 267 HADRNTFHRPDKNLYNCPGQRLREIFLQKDLQGRFLATKQVSDLSASKYQWA 326
QY 249 EYRISYGRKQSEMDOLASIVNNELHSGNVWLVQI PRLYNYVKEMGIVTSFQNLDDN 308
DB 327 ELRSYIGKSPDEWYKLAWINDVYSSNIRWLIQIPRLFDIFKSNKMKMSQFELNNI 386
QY 309 FVPLFEVITDPASHQPLHFLKQVGLDLVDDESKEPRP--TKHMTPEQWTVNFPAP 366
DB 387 FLPLFEATAPSKHPELHRLQVIGFSDVDDSKPE-NPLFNDVPRPEEWTYENPPY 445
QY 367 SYAYCYANLFTYKLRBSKGMTHTKFRPHAGEAGDVHAAATFLLCNHSIGINLRS 426
DB 446 AYIYIYMYANMTVNLKFRGSRNMTFVLRPHCGEAGVQHLVCGFLMAENISHGLLRKV 505
QY 427 PVLQYLYLQIGLMSPLSNLSFLDYHNRNPTTFQRLNVLSTDDPQLQHLKPEL 486
DB 506 PVLQYLYLQIGLMSPLSNLSFLDYHNRNPTTFQRLNVLSTDDPQLQHLKPEL 565
QY 487 VEYSIAASLWKLSCDLCBIARNVYQSGFSHALKAHMTGKNGYKRGAGNDIHRNTVP 546
DB 566 MEYSIAAQVWKLSCDLCBIARNVYQSGFSHALKAHMTGKNGYKRGAGNDIHRNTVP 625
QY 547 HIRVQREMIRWEMKLVYSDNE 569
DB 626 EIRVAYRYETLLDELNSIFKVNQ 648

RESULT 11
ABG70650
ID ABG70650 standard; Protein; 747 AA.
XX AC
XX ABG70650;
XX DT
XX 12-DEC-2002 (first entry)
XX DE
XX Rabbit adenosine monophosphate deaminase (AMPDA).
XX KW
XX Rabbit; adenosine monophosphate deaminase; AMP deaminase; AMPDA;
KW hydrolysis; 5'AMP; inosine monophosphate; IMP; X-ray diffraction;
KW crystal structure; ischaemia-related disease; Alzheimer's disease;
KW congestive heart failure; peripheral vascular disease; CAD;
KW coronary artery disease; chronic obstructive pulmonary disease;
KW vasotropic; cardiant; cardiovascular; nootropic; neuroprotective;
KW EC 3.5.4.6; protein co-ordinate data; enzyme.
XX OS
XX Oryctolagus cuniculus.
XX PN
XX EP1215587-A2.
XX PD
XX 19-JUN-2002.
XX PF
XX 29-NOV-2001; 2001EP-0309996.
XX PX
XX 13-DEC-2000; 2000GB-0030424.
XX PA
XX (PFIZ) PFIZER LTD.
XX PA
XX (PFIZ) PFIZER INC.
XX PI
XX Bazin RJ, McDonald GA, Phillips C;
XX WPI; 2002-521745/56.
XX DR
XX N-PSDB; ABS54378.
XX PT
XX New rabbit AMP deaminase crystal useful for identifying AMP deaminase
PT inhibitors that can treat e.g. congestive heart failure and Alzheimer's
XX Claim 36; Page 449-450; 459pp; English.
XX The present invention relates to the crystal structure of rabbit
XX adenosine monophosphate (AMP) deaminase (AMPDA), which is of
XX tetragonal form. AMPDA catalyses the hydrolysis of 5'AMP to form
XX inosine monophosphate (IMP) and ammonia. The high resolution
XX structure of AMPDA is obtained by X-ray diffraction. The atomic
XX co-ordinates of the crystal can be used to derive the 3-dimensional
XX structure of AMPDA, which can be used to evaluate binding interactions
XX of chemical compounds with AMPDA and design associating compounds
XX e.g. AMPDA inhibitors. The selected compound is useful for treating
XX ischaemia-related diseases (e.g. congestive heart failure,
XX peripheral vascular disease, coronary artery disease (CAD)), chronic
XX obstructive pulmonary disease, and Alzheimer's disease. The
XX co-ordinates are also useful for solving a crystal form of a mutant,
XX homologue or co-complex of AMPDA. The present sequence represents
XX rabbit AMPDA.
SQ Sequence 747 AA;
Query Match 48.5%; Score 1502.5; DB 23; Length 747;
Best Local Similarity 49.9%; Pred. No. 3.4e-140;
Matches 286; Conservative 96; Mismatches 168; Indels 23; Gaps 8;
QY 8 WEKEVINPDC-TP--KPNPMPF--TYVPEPKSEHVFQTVGVHIVYADKCTES----- 56
DB 174 WKANESSYVFTPALKKGEDPRTDNLNENLGYHL-KMKDGVVYIYANEAAAGDKPKPL 232
QY 57 IYPVADATFTTDLHYILRVTAAGNTRVTCNRLNLEHKKFKHMLNADREFLAQKTAH 116
DB 233 LYP--NMEEFLDDNMNLLALLAQGPVKTVTHRRLLKFLSKFQVHQLMDELKELKNP 290


```
XX 16-MAY-2002 (first entry)
XX A thaliana AMP deaminase EST encoded fragment.
XX AMP deaminase; adenosine monophosphate; transgenic plant; EST;
KW herbicide resistance; herbicide; inhibitor; expressed sequence tag.
XX Arabidopsis thaliana.
XX WO200206319-A2.
XX 24-JAN-2002.
XX 06-JUL-2001; 2001WO-EP07767.
XX 17-JUL-2000; 2000DE-1035084.
XX (AVET ) AVENTIS CROPS SCIENCE GMBH.
XX Schulz A, Streiber W, Hanke C, Schmidt F, Schubel A;
XX WPI; 2002-195802/25.
XX N-PSDB; AAL45031.
XX New nucleic acid for plant adenosine monophosphate deaminase, useful in
PT screening for herbicides and preparing herbicide-resistant plants -
XX Example 2; Page 23; 51pp; German.
XX The present invention provides the protein and coding sequences of the
CC Arabidopsis thaliana adenosine monophosphate (AMP) deaminase. The coding
CC sequence can be used to transform prokaryotic or eukaryotic cells,
CC especially to produce transgenic plants (e.g. barley, rice, soya etc.)
CC resistant to herbicidal inhibitors of AMP-deaminase, for recombinant
CC production of proteins with AMP-deaminase activity, and to identify
CC related genes in other organisms. The protein can be used for
CC identification and biochemical/structural characterisation of new
CC AMP-deaminase inhibitors and potential herbicides. The present sequence
CC is an EST encoded fragment of the protein of the invention.
XX SQ Sequence 292 AA;
Query Match 43.8%; Score 1357.5; DB 23; Length 292;
Best Local Similarity 84.2%; Pred. No. 2.2e-126;
Matches 245; Conservative 24; Mismatches 21; Indels 1; Gaps 1;
QY 286 PRLNYYKEMGIVTSFQNLNDNIFVPLFEVITDPASHPQLHVLKQVGLDLVDDESKPE 345
DB 1 PRLNYYKONGIVTSFQNLNDNIFPLFEATVPDHPQLHVLKQVGLDLVDDESKPE 60
QY 346 RRPTKHMPTEQWNTNVPNPAFSYAYCYANLFTLNKLRSGMTTIKFRPHAGEAGDVD 405
DB 61 RRPTKHMPTEQWNTNVPNPAFSYAYCYANLFTLNKLRSGMTTIKFRPHAGEAGDID 120
QY 406 HLAATFLLCNISHGINLRKSPVLYLYLGOIGLWSPNSLSFLDYHRNPFPTFFQF 465
DB 121 HLAATFLTCSHAGINLRKSPVLYLYLGOIGLWSPNSLSFLDYHRNPFPTFFFLR 180
QY 466 GLNVLSTDDPLQIHLTKPELVEEYSTAAASLWKLSSCDLCEIARNYSYQGFSGHALKAHW 525
DB 181 GLNVLSTDDPLQIHLTKPELVEEYSTAAASVWKLSCADLCEIARNYSYQGFSGHALKSHW 240
QY 526 IGRNYFKRGAGNDIHTNTVPHIRVQPREMWRNEMKLVYSDNEILLIPDEL 576
DB 241 IGRDYKRGPGDNDIHTNTVPHIRVQPREMWRNEMKLVYSDNEILLIPDEL 576
RESULT 14
AAG70751
ID AAG70751 standard; Protein; 888 AA.
XX
AC AAG70751;
```

```
XX 27-JUL-2001 (first entry)
XX S cerevisiae apoptosis associated protein YUL035C.
XX Yeast; fungus; apoptosis; infection; proliferative disease;
KW vaccine; autoimmune disease; ischaemia; neurodegeneration.
XX Saccharomyces cerevisiae.
XX WO200102550-A2.
XX 11-JAN-2001.
XX 03-JUL-2000; 2000WO-BE00077.
XX 01-JUL-1999; 99EP-0870141.
XX (JANC ) JANSSEN PHARM NV.
XX Contreras RH, De Backer MD, Luyten WHML, Malcorps IKL;
PI Nellissen BJM, Reekmans RJ;
XX WPI; 2001-367042/38.
XX N-PSDB; AAH29787.
XX Yeast and fungal nucleic acids encoding proteins involved in a pathway
PT leading to programmed cell death, useful for treating proliferative
PT disorders, yeast and fungal infections, or for preventing apoptosis in
PT certain diseases -
XX Claim 1; Fig 1; 218pp; English.
XX The present invention provides the protein and coding sequences of a
CC number of apoptosis associated proteins from the yeast Saccharomyces
CC cerevisiae and the fungus Candida albicans. These can be used to identify
CC treatments for fungal and yeast infections, for proliferative diseases
CC and for apoptosis related diseases such as autoimmune diseases, ischaemia
CC and neurodegeneration. The present sequence is one of the S. cerevisiae
CC proteins of the invention.
XX SQ Sequence 888 AA;
Query Match 20.2%; Score 627; DB 22; Length 888;
Best Local Similarity 30.1%; Pred. No. 8.2e-53;
Matches 194; Conservative 94; Mismatches 207; Indels 126; Gaps 21;
QY 66 PFTDLHYLRVTAAGNTRTVCHNRLNLEHKKFHLMLNADREFLAQKTAHPDRFVNVVK 125
DB 281 FRDDPAYITELIQSHKFNEVSRKLSYLLDKFELFYLNKSKKEILANKVNPYRDFVNSRK 340
QY 126 VDTVHHSACMOKHLLRIKSKRKEPDEVIFRD---GTYMTLKEVF----- 171
DB 341 VDRDLSLSGCISQRLSEYIWEKINLEPERIV-YQDPETSRKLSURDIFQCGSSNDQPI 399
QY 172 -----ESDLATGYDLNVDLLDVHADKSTFHRFDKENKYNPCGQRLREIFLKQDN 222
DB 400 ATGLKLIDDEFID---WYENIYLDIYHLPNPKVAKLVGKEMRF-----YLLAKVFLFDN 451
QY 223 LIQGRFLAEL-TKQVPSDLSASKYQMAEYRIS--IYGRKQSEWDQLASIVNNELHSGNV 279
DB 452 FIEGEYLABIFIKYVHILEKSKYQLAQVSNVQFYSSGEDWYKFSOWLLRWKLVSINI 511
QY 280 VMLVQIPRLY-NVYKEMGIVTSFQNLNDNIFVPLF-----EVTIDPASHPQ---LHVELK 330
DB 512 RWNQIARIFPKLFKE-NVYSNFBQFBLDIFNPLFTLEKEQPIDSSVNTDILGLOFFLS 570
QY 331 QVVGLDLVDDESKPE--RRPTKHMPTPEQWT-NVFNPAFSYAYCYANLFTLNKLRSG 387
DB 571 NVCSMDLVIKESDEYKWKFTDMNCKKFTWAGQDNPTVAHYMYIYKSLAKVNFRLSQN 630
QY 388 GMTTIKFRPHAGE-----AGDVHDLAATFLLCNISHGINLRKSP----- 427
```


Result No.	Score	Query Match	Length	DB	ID	Description
1	156.5	5.0	516	2	US-09-019-201A-3	Sequence 3, Appli
2	113	3.6	352	4	US-09-328-352-6481	Sequence 6481, Appl
3	111.5	3.6	331	4	US-09-252-991A-28467	Sequence 28467, A
4	106.5	3.4	999	3	US-09-101-886B-2	Sequence 2, Appli
5	105	3.4	464	1	US-07-991-867B-2	Sequence 2, Appli
6	105	3.4	464	1	US-08-107-755A-2	Sequence 2, Appli
7	105	3.4	464	2	US-08-544-332-2	Sequence 2, Appli
8	105	3.4	464	4	US-09-370-861A-2	Sequence 2, Appli
9	104.5	3.4	511	2	US-09-019-201A-2	Sequence 2, Appli
10	100	3.2	604	3	US-09-586-935-3	Sequence 3, Appli
11	100	3.2	604	4	US-09-872-861-4	Sequence 4, Appli
12	99.5	3.2	371	1	US-08-487-748A-12	Sequence 12, Appl
13	99.5	3.2	371	3	US-08-480-070C-12	Sequence 12, Appl
14	99.5	3.2	371	3	US-08-629-525-12	Sequence 12, Appl
15	99.5	3.2	371	3	US-08-609-583A-12	Sequence 12, Appl
16	99.5	3.2	371	3	US-08-937-399-12	Sequence 12, Appl
17	99.5	3.2	371	4	US-09-310-367-12	Sequence 12, Appl
18	99.5	3.2	371	4	US-09-032-337-12	Sequence 12, Appl
19	99.5	3.2	371	4	US-09-464-231-12	Sequence 12, Appl
20	99.5	3.2	523	4	US-09-555-889A-2	Sequence 2, Appli
21	99	3.2	1651	1	US-08-447-411-2	Sequence 2, Appli
22	98	3.2	921	1	US-07-872-644-39	Sequence 39, Appl
23	98	3.2	921	1	US-08-297-494-39	Sequence 39, Appl
24	98	3.2	921	1	US-08-297-510-39	Sequence 39, Appl
25	98	3.2	921	1	US-08-479-532-39	Sequence 39, Appl
26	98	3.2	921	1	US-08-455-526-39	Sequence 39, Appl
27	98	3.2	921	1	US-08-455-525-39	Sequence 39, Appl

Qy	81	NRTVTCHNRNLNLEHKFPFHLMLNADREFLA---OKTAPH-----RDFVN 122
Db	143	NL-WVCQGK-----GDKEVIGMRFSEKTKPDVATQADCTWELLSKVRELHG 186
Qy	123	VVKVDTHVHSSACMQKHLLFIKSRLKRPDEVVIFRDGTYYMTLKE--VPFESLDLTGYD 180
Db	187	ADKVDTYL-----REHLTYPYTVKFJLDNNEAQEF--GSIIPALJDGLLFYAPSWADYY 237
Qy	181	LNVDDLVDHADSKTFHRFDKFNLYNPGCQSRLREIFLKQDNITQGRFLAEL-TKQVFSD 239
Db	238	YNA-LKEFHADGVQYLEF-----RSTLPILY-----DLEGTSFELODTVRIYKE 280
Qy	240	LSASKYQMAEYRISITYGRKQSEWDQLASWINNELHSGNVVWLVOPLRNIVNYKEMGIYT 299
Db	281	TLDKY-MAEH-IDFIGSK-----IYAP-IINTDKEG-----309
Qy	300	SFQNLLDNIFVPLEFTVIDPASHPOLHVFLKQVVGVLDDDESKEPRERTKHMPTPEQWT 359
Db	310	-----LDNYIKVCVEI---KEKYPDF-----VAGFLVGOEEK--GRPLKDP-----346
Qy	360	NVENPAFSYAYCYANLFTLNKLRESKGM-TTIKFRPHAGEAG-----DVDHLAAETL 412
Db	347	-----IPOLLGMPENIDFYFHAGETNWFGSTVDENLIDAVILL 383
Qy	413	LCNHTISHGINLRKSPVOLYLYLGOGIAMSPLSNSLFL--DYHRNPPTFFORGLNVIS 470
Db	384	GTRKIGHGFALVKHPLVLMQLKERNIAIEVNPI SNQVLQLVADYRNHPCAIFYPADNPYV 443
Qy	471	LSTDPPQLQHLPKEPLVBZEYSIAASLWKLSCD--LCEIARNSVVOGSFSHALK----A 523
Db	444	ISDDPSFWKAT--PLSHDFYFIAFGIASAHSDMRLLKKUALANSINYSSLSPQKRVALA 501
Qy	524	HW 525
Db	502	KW 503

```

RESULT 2
US-09-328-352-6481
; Sequence 6481, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6481
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6481

```

Db	138	LRRACADAKDKGGISQI	IMCFRLHSEEAFTL	172
Qy	317	IDPASHQHLVFKQVGLD	VDDESKPERPTKHMPT	376
Db	173	-----EQALPEKQDII	AVGL--DSSEVGHPPAK	204
Qy	377	LFTLNKLRESKGMTTI	KERPHAGAGVDHLL--	434
Db	205	-----KARQ--EGFLIV	--AHAGEGPEYVWEAL	255
Qy	435	LQOIGLAMPISNNSLFL	--DYHRNPPTFPORGL	492
Db	256	NEKMPLTVCPLSNLKL	CVVKKOMKHNIRLLQK	312
Qy	493	AASLWKLSSCDLCEI	ARNSVYQSGFSHALKA	528
Db	313	IOAALDITNDELKOLA	INSFEASISDEEKOK	348

```

RESULT 3
US-09-252-991A-28467
; Sequence 28467, Application US/09252991A
; Patent NO. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28467
; LENGTH: 321
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28467

```

```

RESULT 4
US-09-101-886B-2
; Sequence 2, Application US/09101886B
; Patent No. 6197507
; GENERAL INFORMATION:
; APPLICANT: BERG, THOMAS
; APPLICANT: TOLLERSRUD, OLE K
; APPLICANT: NILSEN, OIVIND
; TITLE OF INVENTION: GENETIC TEST FOR ALPHA-MANNOSIDOSIS
; NUMBER OF SEQUENCES: 104

```

Qy	426	SPVLQYLYLGGIGLAMSPLSNNSLFLDYHRNPFPTFFQRGLNVLSLSTDDPLQIHLTKPEP	489
Db	644	QASGAYIF-----RPNQNKPLFVSHW-----AQTHLVKAS	673
Qy	486	LVEEYSIAASLM	497
Db	674	LVOEVHONFSAW	685
 RESULT 5 US-07-991-867B-2 ; Sequence 2, Application US/07991867B ; Patent No. 5476781 ; GENERAL INFORMATION: ; APPLICANT: Moyer, Richard W. ; APPLICANT: Hall, Richard L. ; APPLICANT: Gruidl, Michael E. ; TITLE OF INVENTION: No. 5476781el Entomopoxvirus Expression System ; NUMBER OF SEQUENCES: 66 ; CORRESPONDENCE ADDRESS: ; ADDRESSEE: David R. Saliwanchik ; STREET: 2421 N.W. 41st Street, Suite A-1 ; CITY: Gainesville ; STATE: FL ; COUNTRY: USA ; ZIP: 32606 ; COMPUTER READABLE FORM: ; MEDIUM TYPE: Floppy disk ; COMPUTER: IBM PC compatible ; OPERATING SYSTEM: PC-DOS/MS-DOS ; SOFTWARE: Patent In Release #1.0, Version #1.25 ; CURRENT APPLICATION DATA: ; APPLICATION NUMBER: US/07/991,867B ; FILING DATE: 12-DEC-1992 ; CLASSIFICATION: 435 ; PRIOR APPLICATION DATA: ; APPLICATION NUMBER: WO 92/14818 ; FILING DATE: 12-FEB-1992 ; PRIOR APPLICATION DATA: ; APPLICATION NUMBER: US 07/827,685 ; FILING DATE: 30-JAN-1992 ; PRIOR APPLICATION DATA: ; APPLICATION NUMBER: US 07/657,594 ; FILING DATE: 19-FEB-1991 ; ATTORNEY/AGENT INFORMATION: ; NAME: Saliwanchik, David R. ; REGISTRATION NUMBER: 31,794 ; REFERENCE/DOCKET NUMBER: UF114.C3 ; TELECOMMUNICATION INFORMATION: ; TELEPHONE: 904-375-8100 ; TELEFAX: 904-372-5800 ; INFORMATION FOR SEQ ID NO: 2: ; SEQUENCE CHARACTERISTICS: ; LENGTH: 464 amino acids ; TYPE: amino acid ; TOPOLOGY: linear ; MOLECULE TYPE: protein ; US-07-991-867B-2			
Query Match 3.4%; Score 105; DB 1; Length 464; Best Local Similarity 22.0%; Pred. No. 0.028; Matches 72; Conservative 50; Mismatches 123; Indels 82; Gaps 17			
Qy	25	PFTYVPKPS---EHVFQIVDGVHVIYADKCTESIYPVADATTFTDLHYILRVTAAGN	81
Db	167	PLGFNPCKPKPRKKDKST-----WLSSGDYINCIIYPLTMTINDYDFHLILFEKTOKN	220
Qy	82	TRTVCHN-RLNLELHKFKPHMLNADREFLAQKTAPHDRFDYFNVRKVDTVHHSCMNOKH	140
Db	221	IATVASSMCYKLEDEVKFFLMDNDKKRFPM-----FPIYN-----DIFTCCVIDKH	267
Qy	141	LIR-----FIKSRLKEPDEVIFPDGYTMTLK--EFVESLDLTGYD-LNYVDL	185

[illegible]

Db 268 FDKKAAAYFFNSSGYPBELIKQNKYMFIESD---MTIKSHKYNSTNTNVAIYIDV 324
QY 186 LDVHAD---KSTFHRFDKFNLYN--PCQSRLREIFLKQDNLIQGRFLAELTKOVFS- 238
Db 325 LSEYLDIFKNVNYFFNTFELQYDSDPCGMFNIIFLYIVVFNKSKF--EFKLYYSM 382
QY 239 ----DLASAKYQMA-----EYR--ISYGRKQSEWDQLASWVNNELHSGNVV 280
Db 383 SFIGDLLASSYRGALFISRYDINSIDFKNLTLEIFNKNKFMELIDMYKKS----- 435
QY 281 WLVOIPRLYNYKEMGIVTSFQNLN 307
Db 436 -----NRIMNVCSK--IKNDYDSYIDN 455

RESULT 6

US-08-107-755A-2
; Sequence 2, Application US/08107755A
; Patent No. 5721352
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Hall, Richard L.
; APPLICANT: Gruidl, Michael E.
; TITLE OF INVENTION: No. 5721352el Entomopoxvirus Expression System
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: U.S.A.
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/107,755A
; FILING DATE: 19-AUG-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: US 07/827,658
; FILING DATE: 30-JAN-1992
; PRIOR APPLICATION NUMBER: US 07/657,584
; FILING DATE: 19-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF114.C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (904) 375-8100
; TELEFAX: (904) 372-5800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 464 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-107-755A-2

Query Match 3.4%; Score 105; DB 1; Length 464;
Best Local Similarity 22.0%; Pred. No. 0.028;
Matches 72; Conservative 50; Mismatches 123; Indels 82; Gaps 17;
QY 25 PFTYVPEPKS---EHVFQTVGVHIVYADKCTESIYPVADATFTFDLHYLRVTAAGN 81
Db 167 PLGFNPKPXPYKDKST-----WLSGDIYNCIYPLTMTINDYDYEHLILFEKTDKN 220
QY 82 TRTVCHN-RUNLLEHKFKHMLNADREFLAQKTAHPDRFYNVRKVDTHVHHSACNNQGH 140
Db 221 IATVASSMRCYKLEDRVKFLMNDKKRFM-----FPIIYN-----DHFTCCVIDKH 267

QY 141 LLR-----FIKSLRKEPDEVVIFRDGTYMTLK--EVFESLDLTGYD-LNVDL 185
Db 268 FDKKAAAYFFNSSGYPBELIKQNKYMFIESD---MTIKSHKYNSTNTNVAIYIDV 324
QY 186 LDVHAD---KSTFHRFDKFNLYN--PCQSRLREIFLKQDNLIQGRFLAELTKOVFS- 238
Db 325 LSEYLDIFKNVNYFFNTFELQYDSDPCGMFNIIFLYIVVFNKSKF--EFKLYYSM 382
QY 239 ----DLASAKYQMA-----EYR--ISYGRKQSEWDQLASWVNNELHSGNVV 280
Db 383 SFIGDLLASSYRGALFISRYDINSIDFKNLTLEIFNKNKFMELIDMYKKS----- 435
QY 281 WLVOIPRLYNYKEMGIVTSFQNLN 307
Db 436 -----NRIMNVCSK--IKNDYDSYIDN 455

RESULT 7

US-08-544-332-2
; Sequence 2, Application US/08544332
; Patent No. 5935777
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Hall, Richard L.
; APPLICANT: Gruidl, Michael E.
; TITLE OF INVENTION: No. 5935777el Entomopoxvirus Expression System
; NUMBER OF SEQUENCES: 77
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Gerard H. Bencen
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/544,332
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/991,867
; FILING DATE: 07-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/107,755
; FILING DATE: 19-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 92/14818
; FILING DATE: 12-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,685
; FILING DATE: 30-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/657,584
; FILING DATE: 19-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Bencen, Gerard H.
; REGISTRATION NUMBER: 35,746
; REFERENCE/DOCKET NUMBER: UF114.C4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 464 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-544-332-2

NAME: Coruzzi, Laura A.
 REGISTRATION NUMBER: 30,742
 REFERENCE/DOCKET NUMBER: 7853-081
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-790-9090
 TELEFAX: 212-869-8864
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 371 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 US-08-829-525-12

Query Match 3.2%; Score 99.5; DB 3; Length 371;
 Best Local Similarity 21.8%; Pred. No. 0.075;
 Matches 96; Conservative 49; Mismatches 149; Indels 147; Gaps 22;

QY 165 MTLK---EVFESLDLTGYDINVDLDDVHAD-----KSTFHRFD-KFNLYK-NPCQSRRL 213
 DB 1 MTLTAHLSYFLVLLLAGQGLSDSLTKDAGPRPLEKEVFKLFQIRFNRSYWNPAYETRR 60
 QY 214 REIFLKQDNLIOGRFL-----AELTKQVFSLSASKYQMAEYRISYGRKQS--EWDQ 264
 DB 61 LSIF--AHNLAQAORLQOEDLGTAEFGTPEFSDLTETEEFG-----QLYQERSPERTN 112
 QY 265 LASWIVNNELHSGNVVWLQVIRLYNVYKEMGIVTSFQ-----LLDNIFVPLF 313
 DB 113 MTKKVESN-----TWGESVPTCDWRKAKNIISVKNQGSCKCWMAMAADNI-QALW 164
 QY 314 EVTIDPASHQPLHVFVKQVGLDLDVDESKPERRPTKHMPTPEQWTVNFNPAFSYAYYC 373
 DB 165 RIK-----HQQFVDVSQELLD-----CERCNGCNGGFVWDAY-- 198
 QY 374 YANLFTANKLRESKGMWTKIKRPHAGEAGVDVHDLAATFLCHNISHGINLRKSPVLQYLY 433
 DB 199 ---LTVLN---NSGLASEKDYFQGDGRKPHRCLAKKY-----KKVAMIQ--- 236
 QY 434 YLQIGLAMSPLSNNSLFLDYHRNPPFTFFQRLNVSLSLTDPLQIHLTKPLVEEYSIA 493
 DB 274 VIKATPSSCDPRQV-DHSVLLVGFGEKEGMGTGVLSHSRKRHRSSPYWLKNSWGAHW 332
 QY 526 IGKNYFKRGPGAGNDIHRNTP 546
 DB 333 GEKGYFRLYRGNNTCGVTKYP 353

RESULT 15

US-08-609-583A-12
 Sequence 12, Application US/08609583A
 Patent No. 6204371
 GENERAL INFORMATION:
 APPLICANT: Levinson, Douglas A.
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
 TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
 NUMBER OF SEQUENCES: 37
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pennie & Edmonds
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10036/2711
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq Version 2.0

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/609,583A
 FILING DATE: 01-MAR-1996
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/487,748
 FILING DATE: 07-JUN-1995
 APPLICATION NUMBER: US 08/398,633
 FILING DATE: 03-MAR-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Coruzzi, Laura A.
 REGISTRATION NUMBER: 30,742
 REFERENCE/DOCKET NUMBER: 7853-048
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-790-9090
 TELEFAX: 212-869-8864
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 371 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 US-08-609-583A-12

Query Match 3.2%; Score 99.5; DB 3; Length 371;
 Best Local Similarity 21.8%; Pred. No. 0.075;
 Matches 96; Conservative 49; Mismatches 149; Indels 147; Gaps 22;

QY 165 MTLK---EVFESLDLTGYDINVDLDDVHAD-----KSTFHRFD-KFNLYK-NPCQSRRL 213
 DB 1 MTLTAHLSYFLVLLLAGQGLSDSLTKDAGPRPLEKEVFKLFQIRFNRSYWNPAYETRR 60
 QY 214 REIFLKQDNLIOGRFL-----AELTKQVFSLSASKYQMAEYRISYGRKQS--EWDQ 264
 DB 61 LSIF--AHNLAQAORLQOEDLGTAEFGTPEFSDLTETEEFG-----QLYQERSPERTN 112
 QY 265 LASWIVNNELHSGNVVWLQVIRLYNVYKEMGIVTSFQ-----LLDNIFVPLF 313
 DB 113 MTKKVESN-----TWGESVPTCDWRKAKNIISVKNQGSCKCWMAMAADNI-QALW 164
 QY 314 EVTIDPASHQPLHVFVKQVGLDLDVDESKPERRPTKHMPTPEQWTVNFNPAFSYAYYC 373
 DB 165 RIK-----HQQFVDVSQELLD-----CERCNGCNGGFVWDAY-- 198
 QY 374 YANLFTANKLRESKGMWTKIKRPHAGEAGVDVHDLAATFLCHNISHGINLRKSPVLQYLY 433
 DB 199 ---LTVLN---NSGLASEKDYFQGDGRKPHRCLAKKY-----KKVAMIQ--- 236
 QY 434 YLQIGLAMSPLSNNSLFLDYHRNPPFTFFQRLNVSLSLTDPLQIHLTKPLVEEYSIA 493
 DB 237 -----DFTMLSNNEQAIAHY-----LAVHGPITVTINMK-LLOHYQKG 273
 QY 494 ASLWKLSSCDLCEIARNYSYOSGF-----SHALK-----AHW 525
 DB 274 VIKATPSSCDPRQV-DHSVLLVGFGEKEGMGTGVLSHSRKRHRSSPYWLKNSWGAHW 332
 QY 526 IGKNYFKRGPGAGNDIHRNTP 546
 DB 333 GEKGYFRLYRGNNTCGVTKYP 353

Search completed: December 15, 2003, 13:29:13
 Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 13:28:14 ; Search time 36 Seconds
(without alignments)
2986.071 Million cell updates/sec

Title: US-10-019-633-2

Perfect score: 3100

Sequence: 1 PRRVAPWEKEVINDPCTPK.....NEMKLVSYSDNEILIPDEL 578

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1534	49.5	801	12	US-10-205-219-187
2	838	27.0	446	15	US-10-012-140-19
3	191.5	6.2	160	15	US-10-106-698-5825
4	184	5.9	65	9	US-09-864-761-42950
5	156.5	5.0	516	12	US-10-423-907-3
6	147	4.7	358	15	US-10-156-761-10132
7	111	3.6	198	12	US-10-021-571-8
8	107	3.5	342	15	US-10-156-761-12440
9	106.5	3.4	1032	11	US-09-954-987B-192
10	106.5	3.4	1032	12	US-10-272-502A-31
11	104.5	3.4	333	15	US-10-156-761-13109
12	104.5	3.4	511	12	US-10-423-907-2
13	103	3.3	1771	12	US-10-117-229-4
14	103	3.3	1771	12	US-10-117-229-11
15	102	3.3	654	12	US-10-094-749-1854

16	102	3.3	1715	12	US-10-021-571-4	Sequence 4, Appli
17	102	3.3	1715	12	US-10-117-229-2	Sequence 2, Appli
18	101.5	3.3	1137	12	US-10-259-165-330	Sequence 330, App
19	101	3.3	796	11	US-09-950-041-41	Sequence 41, Appl
20	99.5	3.2	371	12	US-10-252-131-12	Sequence 12, Appl
21	99.5	3.2	371	15	US-10-004-633-12	Sequence 12, Appl
22	99.5	3.2	622	11	US-09-252-088-9	Sequence 9, Appli
23	99	3.2	466	15	US-10-081-872-198	Sequence 198, App
24	99	3.2	682	12	US-10-032-585-7047	Sequence 7047, Ap
25	99	3.2	871	12	US-10-032-585-7050	Sequence 7050, Ap
26	98.5	3.2	1032	15	US-10-088-567-4	Sequence 4, Appli
27	98.5	3.2	1184	12	US-10-117-229-3	Sequence 3, Appli
28	98.5	3.2	1184	12	US-10-117-229-9	Sequence 9, Appli
29	98	3.2	892	10	US-09-895-913A-198	Sequence 198, App
30	98	3.2	921	10	US-09-883-825-39	Sequence 39, Appl
31	98	3.2	921	14	US-10-094-989-5	Sequence 5, Appli
32	98	3.2	942	10	US-09-883-825-43	Sequence 4, Appli
33	97.5	3.1	2471	14	US-10-116-048-4	Sequence 4, Appli
34	96.5	3.1	1032	11	US-09-954-887B-3	Sequence 3, Appli
35	96.5	3.1	1032	12	US-10-272-502A-9	Sequence 9, Appli
36	96.5	3.1	1032	12	US-10-265-072-8	Sequence 8, Appli
37	96.5	3.1	1987	12	US-10-032-585-7518	Sequence 7518, Ap
38	94.5	3.0	904	11	US-09-950-041-6	Sequence 6, Appli
39	94.5	3.0	904	12	US-10-265-072-2	Sequence 2, Appli
40	94.5	3.0	1032	11	US-09-842-758-34	Sequence 34, Appli
41	94.5	3.0	1073	9	US-09-819-249-2	Sequence 2, Appli
42	94	3.0	1148	14	US-10-108-605-67	Sequence 67, Appl
43	94	3.0	1148	14	US-10-108-605-173	Sequence 173, App
44	94	3.0	1148	14	US-10-108-605-275	Sequence 275, App
45	93.5	3.0	1073	15	US-10-157-031-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1

US-10-205-219-187
Sequence 187, Application US/10205219
Publication No. US20030138803A1
GENERAL INFORMATION:
APPLICANT: Warner-Lambert Company
APPLICANT: Lee, Kevin
APPLICANT: Dixon, Alistair
APPLICANT: Brooksbank, Robert
APPLICANT: Pinnock, Robert
TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
FILE REFERENCE: WL-A-018200
CURRENT APPLICATION NUMBER: US/10/205,219
CURRENT FILING DATE: 2002-07-24
PRIOR APPLICATION NUMBER: GB 0118354.0
PRIOR FILING DATE: 2001-07-27
NUMBER OF SEQ ID NOS: 197
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 187
LENGTH: 801
TYPE: PRT
ORGANISM: Rattus norvegicus
FEATURE:
OTHER INFORMATION: AMP deaminase isoform C
US-10-205-219-187

Query Match 49.5%; Score 1534; DB 12; Length 801;
Best Local Similarity 51.8%; Pred. No. 1.7e+140;
Matches 288; Conservative 90; Mismatches 164; Indels 14; Gaps 4;
QY 19 PKPNPNPTTYPEPKS-EHVEQTVDGVHVVADKCTE-----SIYPVADATFTFDLHYI 73
Db 236 PLPQEDPYCLDDAPPENLGLVVMQGVLFVYDNTMLERQPHSLPYPDLETIVDMSHI 295
QY 74 LRVTAGNTRVCNRLNLEHKKFKHMLNADREFLAQKTAHRDFNVNPKVTHVHS 133
Db 296 LALITDGTPTKTYCHRRLNFLSKFSLHMLNEMSEFKELKSNPHRDFNVNPKVTHVHAA 355

APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Aeonica-X-1
CURRENT APPLICATION NUMBER: US 09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 42950
LENGTH: 65
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL096773.6
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.99
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.92
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.9
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
OTHER INFORMATION: EST HUMAN HIT: AW342074.1, EVALUATE 6.00e-32
OTHER INFORMATION: SWISSPROT HIT: P23109, EVALUATE 8.00e-33
US-09-864-761-42950

Query Match 5.9%; Score 184; DB 9; Length 65;
Best Local Similarity 56.9%; Pred. No. 2.7e-10;
Matches 37; Conservative: 7; Mismatches 21; Indels 0; Gaps 0;

QY 126 VDTTHVHSAACMNQKHLRFKSKLRKEPDEVVIFRDGTYMTLKEVFESLDLTGYDLNVDL 185
DB 1 VDTTHVHSAACMNQKHLRFKSKLRKEPDEVVIFRDGTYMTLKEVFESLDLTGYDLNVDL 185
QY 186 LDVHA 190
DB 61 LDVHA 65

RESULT 5

US-10-423-907-3
Sequence 3, Application US/10423907
Publication No. US20030186383A1
GENERAL INFORMATION:
APPLICANT: Fan et al.
TITLE OF INVENTION: Dendritic Cell-Derived Growth Factor
FILE REFERENCE: PF346D1C1
CURRENT APPLICATION NUMBER: US/10/423,907
CURRENT FILING DATE: 2003-04-28
PRIOR APPLICATION NUMBER: 09/360,149
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 09/019,201
PRIOR FILING DATE: 1998-02-05
PRIOR APPLICATION NUMBER: US 60/038,829
PRIOR FILING DATE: 1997-02-06
NUMBER OF SEQ ID NOS: 11
SOFTWARE: Patent in version 3.2
SEQ ID NO 3
LENGTH: 516
TYPE: PRT
ORGANISM: Sarcophaga perigrina
US-10-423-907-3

Query Match 5.0%; Score 156.5; DB 12; Length 516;
Best Local Similarity 22.3%; Pred. No. 3.2e-06;
Matches 121; Conservative 65; Mismatches 189; Indels 167; Gaps 30;

QY 25 PFTVPEKSEHVFTQVDGV-----IHVADKDCSTESYFVADATFTFFDLHYILRVTAAG 80
DB 88 PHLFTP---SCHLFEVLDDIKQSPFKYISSMPKGAVLHAHDALCSTD--FLIRLTYRD 142
QY 81 NTRTVCHNRLNLEHKKFHLMLNADREFLA---OKTAPH-----RDFTN 122
DB 143 NL-WVCOGK-----GDKEVIGMRFSKRPDVATQADCTWELLSSKVRLEHG 186
QY 123 VRKVDTHVHSAACMNQKHLRFKSKLRKEPDEVVIFRDGTYMTLKE--VFESLDLTGYD 180
DB 187 ADKVDTYL-----REHLYPTVKFLDNNNEQF--GSIFALLDGLLFYAPSWADYY 237
QY 181 LNVDLLDVHADKSTFHRPDKFNKYNPGQSRLEIFLKQDNLIQGRFLAEL-TKQVPSD 239
DB 238 YNA-LKEFHADGVQYLEF-----RSTLPILY-----DLEGTSTELDTVRIYKE 280
QY 240 LSASKYQMAEYRISYIGRKQSEWDLASWIVNNELHSGNVVWLVQIPLRYNVYKMGIVT 299
DB 281 -TLDKY-MAEH-IDFIGSK-----LIYAP-IRNTDKEG-----309
QY 300 SFQNLNDNIFVPLPEVTIDPASHPLQHLVFLKQVVGLDLVDDSKPERPRTKHMPTPEQWT 359
DB 310 -----LDNYIKVCVEI---KEKYPDF-----VAGFLVLVQEEK--GRPLKDF-----346
QY 360 NVFNPAFSYAYCYANLFTLNKLRKSGM-TTIKFRPHAGEAG-----DVDHLAATFL 412
DB 347 -----IPQLLGMPEIDFVFHAGETWNFGSTVDENLIDAVLL 383
QY 413 LCHNISHGINLRKSPVQLYLYLQIGLAMSPLSNNSLFL--DYHRPPFTFFQRLNVS 470
DB 384 GTRKRGHGFALVGHPLVLQMLKERNIAIEVNPISNQVLQVLVADYRNHPCAYFFADNPVW 443
QY 471 LSTDDPLQIHLTKPELVEEYSAASLWKLSSCD---LCEIARNVSVOGSGFSHALK---A 523
DB 444 ISSDDPSFWKAT--PLSHDFYIAFLGIASHSDRLKLLKALNSINYSLSPEOKRVALA 501
QY 524 HW 525
DB 502 KW 503

RESULT 6
US-10-156-761-10132

; Sequence 10132, Application US/10156761
; Publication No. US20030119018A1

; GENERAL INFORMATION:

; APPLICANT: OMURA, SATOSHI

; APPLICANT: IKEDA, HARUO

; APPLICANT: ISHIKAWA, JUN

; APPLICANT: HORIKAWA, HIROSHI

; APPLICANT: SHIBA, TADAYOSHI

; APPLICANT: SAKAKI, YOSHIYUKI

; APPLICANT: HATTORI, MASAHIRA

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

; FILE REFERENCE: 249-262

; CURRENT APPLICATION NUMBER: US/10/156,761

; CURRENT FILING DATE: 2002-05-29

; PRIOR APPLICATION NUMBER: JP 2001-204089

; PRIOR FILING DATE: 2001-05-30

; PRIOR APPLICATION NUMBER: JP 2001-272697

; PRIOR FILING DATE: 2001-08-02

; NUMBER OF SEQ ID NOS: 15109

; SEQ ID NO 10132

; LENGTH: 358

; TYPE: PRT

; ORGANISM: Streptomyces avermitilis

US-10-156-761-10132

Query Match 4.7%; Score 147; DB 15; Length 358;

Best Local Similarity 23.6%; Pred. No. 1.5e-05;

Matches 100; Conservative 51; Mismatches 156; Indels 116; Gaps 19;

QY 117 HRDFYNNVRKVDTHVHSAACNQHLLRFKSKLRKEPDEVIFRDGTYMTLKEVFESLDL 176

DB 20 HSIAGLPKRAELHVVHVSASP-----RIVSELAARHPDSSV-----PTDPEAL----- 63

QY 177 TGYDLNVLDLVHADKSTHRRPKFNLYNCPGQSLRLEIFLKQDNLIQ-GRFLABLTQ 235

DB 64 -----ADYFTTDFAHF-----IKVYLSVVDLIRTPEDVRLLTYE 98

QY 236 VFSDLSASKYQMAEYRISYGRKQSEMDQLASWIV-----NNELHSGNVV-WLVQIPRL 288

DB 99 VARELARQQRYAELTYTFSSSTRGIDERA FMDATEDARKSAEAFGVVLRWCFDIP-- 156

QY 289 YNIVKEMGIVTSQNLLNIFVPLFEVTTDPASHPQLHVFLKQVGLDLVDDESKPERRP 348

DB 157 -----GEAGLESA-----EETVRLAT-----DORLREGLV 182

QY 349 TKMPTPEQWTVNFPNPAFSYAYVCYANLFTLNKLRSGMTIKRPHAGEAGVDHL- 407

DB 183 SFLGGPE--IGVPRQPKYFDRAIA-----AGLRSV---PHAGETTGPTVM 226

QY 408 -AATFLCHNISHGINLRKSPVLQYLYLQ--IGLAMSPLSNLSL-----FLDYHRNPP 460

DB 227 DAUTDLRAERIGHGTSSAQPKL--LAHLAEHRIPLEVCTPSNIATRAVRLDEH--PVK 282

QY 461 TFFQGINVSLSDDDPLQLHLTKPLVEEYSIAASLWKLSSCDLCBIARNVYVQGFSHA 520

DB 283 EFVRAGVVTVINSDDP---PMFGTDLNNEVIAIARLLDLDERGLAGLAKNSVEASFLDAA 339

QY 521 LKA 523

DB 340 GKA 342

RESULT 7

US-10-021-571-8

; Sequence 8, Application US/10021571

; Publication No. US20030116056A1

; GENERAL INFORMATION:

; APPLICANT: CHAO, Moses V.

; APPLICANT: KONG, Haeyoung

; TITLE OF INVENTION: A TRANSMEMBRANE PROTEIN AS A DOWNSTREAM TARGET OF NEUROTROPHIN AN

; TITLE OF INVENTION: RECEPTOR TYROSINE KINASES, DNA ENCODING SAME AND MONOCLONAL ANTI

; FILE REFERENCE: CHAO11A

; CURRENT APPLICATION NUMBER: US/10/021,571

; CURRENT FILING DATE: 2001-12-19

; PRIOR APPLICATION NUMBER: 60/256,909

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: Patent in version 3.1

; SEQ ID NO 8

; LENGTH: 1498

; TYPE: PRT

; ORGANISM: Drosophila

US-10-021-571-8

Query Match 3.6%; Score 111; DB 12; Length 1498;

Best Local Similarity 18.2%; Pred. No. 0.47;

Matches 90; Conservative 75; Mismatches 173; Indels 156; Gaps 21;

QY 26 FTYYPEPKSEHVQTVDCGVHIVYADKCTESIYPVADATTEFTDLHYILR----- 75

DB 703 FTFADKEKEHILVALYVIAAVMGTLICTHLHLAKVFSLFTSHIRLVKAAVRSESA 762

QY 76 VTAGNTRTVCHNRLNLE-----HKPKFHLMLNADREFLAQTAP 116

DB 763 LTMLGAEVAVMTDMVKCLDAFTNQOSRLGVVIDALDSCDTERILTLNNAVQTL 819

QY 117 HRDFYNNVRKVDTHV-----HHSACNQHLLRFKSK-----KLKREP 153

DB 820 NRPFVLLISVDPHVIKAAEAANSRRLEFTEGGIGGHDFLRLNLVLPVYLQNSGLRKVRQA 879

QY 154 DEVIFR---DGYTWT-----LKEVPESLDLT-----GYDLNVLD 185

DB 880 MTALLFKSGGGDYQTDGPTLGHVSARRLSNASEIISQEKLRGPARGGGKKLRUSE 939

QY 186 LDVHADKSTHRRPKFNLYNCPGQSLRLEIFLKQDNLI--IQGRFLAELTKQVFS 240

DB 940 SVASGTGNLIR-----LGQNPQTVLDLSRVLVTDYSDVNPMSRMLMNVIIYITVRL 994

QY 241 SASKYQMAEYRISYGRKQSEMDQLASWIV--NNELHSGNVVWLVOIPRLYNNYK-EMGI 297

DB 995 KAFQIEFSWYRLSSWNLTEQWPLRASMIVLHHQDFMDSNADESVSLSQSVYEKLPRKLA 1054

QY 298 VTSFQNLID-----NIFVPLFEVTTDPASHPQLHVFLKQV 333

DB 1055 LREAPLLELDRDERKLDALQLHKSDDLVLVADRLFLP-FTINLDP-----YLRKVL 1105

QY 334 GLD--LVDE-----SKPE-----RRPTKHMPTPEQWTVNFPNPAFSYAYVCYANL 377

DB 1106 KEDQOTIEDGSLVQARPSVSNTRMQFPATTVYVPSQ-----AIPPYQMFQNE 1155

QY 378 FTLNKLRESKGMT 391

DB 1156 YPANELR-SRNLST 1168

RESULT 8

US-10-156-761-12440

; Sequence 12440, Application US/10156761

; Publication No. US20030119018A1

; GENERAL INFORMATION:

; APPLICANT: OMURA, SATOSHI

; APPLICANT: IKEDA, HARUO

; APPLICANT: ISHIKAWA, JUN

; APPLICANT: HORIKAWA, HIROSHI

; APPLICANT: SHIBA, TADAYOSHI

; APPLICANT: SAKAKI, YOSHIYUKI

; APPLICANT: HATTORI, MASAHIRA

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

; FILE REFERENCE: 249-262

; CURRENT APPLICATION NUMBER: US/10/156,761

; CURRENT FILING DATE: 2002-05-29

; PRIOR APPLICATION NUMBER: JP 2001-204089

; PRIOR FILING DATE: 2001-05-30

; PRIOR APPLICATION NUMBER: JP 2001-272697

; PRIOR FILING DATE: 2001-08-02

; NUMBER OF SEQ ID NOS: 15109

SEQ ID NO	12440	12441	12442	12443	12444	12445	12446	12447	12448	12449	12450	12451	12452	12453	12454	12455	12456	12457	12458	12459	12460	12461	12462	12463	12464	12465	12466	12467	12468	12469	12470	12471	12472	12473	12474	12475	12476	12477	12478	12479	12480	12481	12482	12483	12484	12485	12486	12487	12488	12489	12490	12491	12492	12493	12494	12495	12496	12497	12498	12499	12500	12501	12502	12503	12504	12505	12506	12507	12508	12509	12510	12511	12512	12513	12514	12515	12516	12517	12518	12519	12520	12521	12522	12523	12524	12525	12526	12527	12528	12529	12530	12531	12532	12533	12534	12535	12536	12537	12538	12539	12540	12541	12542	12543	12544	12545	12546	12547	12548	12549	12550	12551	12552	12553	12554	12555	12556	12557	12558	12559	12560	12561	12562	12563	12564	12565	12566	12567	12568	12569	12570	12571	12572	12573	12574	12575	12576	12577	12578	12579	12580	12581	12582	12583	12584	12585	12586	12587	12588	12589	12590	12591	12592	12593	12594	12595	12596	12597	12598	12599	12600	12601	12602	12603	12604	12605	12606	12607	12608	12609	12610	12611	12612	12613	12614	12615	12616	12617	12618	12619	12620	12621	12622	12623	12624	12625	12626	12627	12628	12629	12630	12631	12632	12633	12634	12635	12636	12637	12638	12639	12640	12641	12642	12643	12644	12645	12646	12647	12648	12649	12650	12651	12652	12653	12654	12655	12656	12657	12658	12659	12660	12661	12662	12663	12664	12665	12666	12667	12668	12669	12670	12671	12672	12673	12674	12675	12676	12677	12678	12679	12680	12681	12682	12683	12684	12685	12686	12687	12688	12689	12690	12691	12692	12693	12694	12695	12696	12697	12698	12699	12700	12701	12702	12703	12704	12705	12706	12707	12708	12709	12710	12711	12712	12713	12714	12715	12716	12717	12718	12719	12720	12721	12722	12723	12724	12725	12726	12727	12728	12729	12730	12731	12732	12733	12734	12735	12736	12737	12738	12739	12740	12741	12742	12743	12744	12745	12746	12747	12748	12749	12750	12751	12752	12753	12754	12755	12756	12757	12758	12759	12760	12761	12762	12763	12764	12765	12766	12767	12768	12769	12770	12771	12772	12773	12774	12775	12776	12777	12778	12779	12780	12781	12782	12783	12784	12785	12786	12787	12788	12789	12790	12791	12792	12793	12794	12795	12796	12797	12798	12799	12800	12801	12802	12803	12804	12805	12806	12807	12808	12809	128
-----------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-----

Qy	359	-----TNVFN--PASYAYAYCYANLFTLKL--RESKGMTTIKPRPHAGEAGD	403
Db	497	DIACLNLSFNANTQVNGTESSPMHIKYDL--TNNRLDFPDNNAFSL---HDLVLD	551
Qy	404	VDHLAATFLLCNHSIGNLRKSPVLQVLYYLQIGIGLAMSPLSNNSLF-----	451
Db	552	LSHNAHYFSIA--GVTHRLGF-----IQNL-----INLRVLNLSHNGIVTLTTESELKSIS	600
Qy	452	-----LDYHRN-----PPTFF-----QRCLNVSLSTDDPLOT-----HLTKEP	485
Db	601	LKELVFGSNRLDHLWANDGKYWISFKSLQNLRLDLSTYNNLQQLPNCAGFNLPSLQEL	660
Qy	486	LVESYSAASLWLK-----SSCDLCEIARNSVY-----OSGFSHALKAHWIGKNYKPROP	536
Db	661	LISGNKLRFFNTWTLQVFPFHLLDLDSRNELYFLPNCILSKFAHSLTLLLSHNHFSLHPS	720
Qy	537	GNDIHRTNVPHIRVQFREMIRWEMKLV	564
Db	721	GFLSEARNLVHLDLSF-----NTIKMI	742

RESULT 11

US-10-156-761-13109
; Sequence 13109, Application US/10156761
; Publication No. US20030119018A1

; APPLICATION N.: 03200301130138
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI

APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN

APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI

APPLICANT: HATTORI, MASAHIRA
APPLICANT: HATTORI, MASAHIRA
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: SAKAKI, YOSHIYUKI

; REFERENCE: 249-262
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-262

FILE REFERENCE: 24-202
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29

; CURRENT FILING DATE: 2002-03-23
 ; PRIOR APPLICATION NUMBER: JP 2001-204089
 ; PRIOR FILING DATE: 2001-05-30

; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02

```

; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 13109

```

```

; LENGTH: 333
; TYPE: PRT

```

ORGANISM: Streptomyces avermitilis
US-10-156-761-13109

Qy	257	RKSEWOLASWIVNNELHSGNVWLVOIPRLYNVYKEMGIWTSQNLLD--NIFVPL--	312
Db	41	KKAYCFDQLOSFL-----NLY--YELMAVLRTQEDEFEDLANAYPRAA	81
Qy	313	-----PBTITDPASHPOLHVLKQVVG-----LDLVDDSKPER	346
Db	82	AQGVRAHAEIIFDPQNAHARGVAMGTIVSEGLWALGRSTENHGVSTQLMCFURDSESASA	141
Qy	347	RPTKHMPPPE-----QWNTNVPAPSYVAYYCYANFLTANKLRSGKMTTIKFRPH	397
Db	142	LETLEAAKPYLDRIIGIDLSAEVGHPPAKFREVEYEAALGLRRV-----AH	189
Qy	398	AGEAGDVVDHL--AAFTFLCHNISHGINLRKSPVLQVLYYLGIOGLAMSPLSNNSL--F	451
Db	190	AGEGPPPYITTEALDVGERIDAGLRCEWDDALVERLVRVRVPITLCPLSNVRLEAVDV	249
Qy	452	LDYHRNPPPTTFQFGLNVSLSDDPL-----QIHLTKEPLVEEYSAAASLWKLSSCD	503
Db	250	LAEH--PLPAPMLDAGLLCTVNSDDPAYEGGYVADNFHVAVRDAL-----GLDQER	296

Qy	504	LCEIARNSVYQSGF	519
		: : : :	
Dd	297	MRELARNSFVASFLEH	312

RESULT 12

US-10-423-907-2

; Sequence 2, Application US/10423907
; Publication No. US20030186383A1

GENERAL INFORMATION:
APPLICANT: Fan et al.

FILE REFERENCE: PF346D1C1

;; CURRENT APPLICATION NUMBER: US/10/42
;; CURRENT FILING DATE: 2003-04-28

;; PRIOR APPLICATION NUMBER: 09/111,111
;; PRIOR FILING DATE: 1999-07-26

; PRIOR APPLICATION NUMBER: US 09/019,201
; PRIOR FILING DATE: 1998-02-05

; PRIOR APPLICATION NUMBER: US 0
 ;
 ; PRIOR FILING DATE: 1997-02-06

```

; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.2

```

```

; SEQ ID NO 2
; LENGTH: 511

```

```

;
; TYPE: PRT
; ORGANISM: Homo sapiens

```

Query Match 3.4%; Score 104.5; DB 12; Length 511;
Best Local Similarity 19.4%; Pred. No. 0.38;

181	QY	--LNVDLDVHADKSTFHRFDKFNKYNPCQSGRLREI	FLKQDNLIQGRFLAELTKOVF	237
181	Db	SLRNF TLVTOHP	EVITNQNVWSKP	211
238	QY	SDLSASKYQMAEYRISYIGRKQSEWDQLASIVNNELHSGNVMLVQIPRLYVNVYKMG-	296	
212	Db	FTISGLIHYAPVPRDYVFRSQ	EFVEDNVLYMEIRARLLPVYELSGE	258
297	QY	IVTSFONLNDIFV	PLF-EVITDPASHQPLHV	330
259	Db	HHDEWSVKTYQEVAQK-FVETHPEFGIKIYSDHRSKOVAVIAESIRMAWGLRIKPT	317	
331	QY	QVVGCLDVADESKPERRPTKHPPTPEQWTVNFNPAFSYAYCYANLFTLANKLRSEKGT	390	
318	Db	VWAGFDVLUGHED	TGHSRLDYKEAL	341
391	QY	TIKFRP	HAGEAG	436
342	Db	MIPAKGVKLPYPHAGETDQWQTSIDENILDALWLNATTRIGHGFALSKHPAVRTYSMKK	401	
437	QY	QIGLAMSPLNNNSLFL-DYHRNPFPTFFQRLNVSLSTDOP	476	
402	Db	DIPVEPCINOVKLVSDLRNHHVATLMAGHPMWISSDDP	443	

RESIII.T 13

RESULTS 13
US-10-117-229-4

03-10-117-223-1
; Sequence 4, Application US/10117229
; Publication No. US20030190625A1

REGISTRATION NO: 02000019002341
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies, Inc.

; AFFILIANT: ORIGENE TECHNOLOGIES, INC.
 ; TITLE OF INVENTION: Human Kidins220Pc
 ; FILE REFERENCE: 9U 104 R1

FILE REFERENCE: 50 10 10
; CURRENT APPLICATION NUMBER: US/10
; CURRENT FILING DATE: 2002-04-08

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 13:21:23 ; Search time 17 Seconds
(without alignments)
1598.908 Million cell updates/sec

Title: US-10-019-633-2

Perfect score: 3100

Sequence: 1 PRVRVAPWEKVINDECTPK.....NEMKLVYSNDNEILIPDELDL 578

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1680.5	54.2	879	1	Q01433: homo sapien
2	1647	53.1	810	1	P15274: saccharomyc
3	1617	52.2	846	1	P50998: schizosacch
4	1549	50.0	767	1	Q01432: homo sapien
5	1534	49.5	765	1	O09178: rattus norv
6	1531	49.4	766	1	O08739: mus musculu
7	1505	48.5	747	1	P10759: rattus norv
8	1497	48.3	747	1	P23109: homo sapien
9	627	20.2	888	1	P40361: saccharomyc
10	625	20.2	797	1	P38150: saccharomyc
11	350	11.3	88	1	Q02356: rattus norv
12	150	4.8	331	1	Q8e84: shewanella
13	149	4.8	334	1	Q97e1: clostridium
14	138.5	4.5	387	1	O8e737: streptomyce
15	121.5	3.9	333	1	Q8fh99: escherichia
16	121.5	3.9	339	1	P53984: streptomyce
17	120.5	3.9	352	1	Q9c19: lactococcus
18	117.5	3.8	333	1	O8x661: escherichia
19	117.5	3.8	333	1	P22333: escherichia
20	115.5	3.7	3011	1	P87319: schizosacch
21	112	3.6	331	1	Q8d6q8: vibrio vuln
22	112	3.6	1956	1	Q04956: plasmidom
23	111.5	3.6	316	1	Q916y4: pseudomonas
24	109.5	3.5	7073	1	P59641: h replicase
25	109	3.5	343	1	Q91016: streptomyce
26	106.5	3.4	999	1	Q29451: bos taurus
27	106.5	3.4	1032	1	P58682: mus musculu
28	105	3.4	464	1	P29817: ameacta moo
29	105	3.4	510	1	P58780: sus scrofa
30	105	3.4	511	1	Q9nzk5: homo sapien
31	105	3.4	626	1	Q07744: lactococcus
32	105	3.4	626	1	Q09145: lactococcus
33	104.5	3.4	347	1	P53909: saccharomyc

34 103 3.3 345 1 ADD_RALSO
35 103 3.3 2223 1 CCAE_DISOM
36 102 3.3 856 1 CNRB_CANPA
37 101.5 3.3 334 1 ADD1_VIBPA
38 101.5 3.3 741 1 PA24_BRARE
39 101 3.3 796 1 TLR6_HUMAN
40 100 3.2 901 1 GREB_BORBU
41 99.5 3.2 324 1 ADD_RHIME
42 99.5 3.2 344 1 ADD_CAUCR
43 99.5 3.2 371 1 CATW_MOUSE
44 99.5 3.2 525 1 AGSA_APLCA
45 99.5 3.2 731 1 MCCA_SOYBN

Q8xx15 ralstonia s
P56699 discopyge o
P33726 canis famil
Q87cf3 vibrio para
P50392 brachydanio
Q9v2c9 homo sapien
O51157 borrelia bu
Q92t48 rhizobium m
Q9a3m3 caulobacter
P56203 mus musculu
P15287 aplysia cal
Q42777 glycine max

ALIGNMENTS

RESULT 1
AMD2_HUMAN
ID AMD2_HUMAN STANDARD; PRT; 879 AA.
AC Q01433; Q14856; Q14857; Q16686; Q16687; Q16688; Q16729; Q9UDX8;
AC Q9UDX9; Q9UMU4;
DT 01-JUL-1993 (Rel. 26, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE AMP deaminase 2 (EC 3.5.4.6) (AMP deaminase isoform L).
GN AMPD2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=93054531; PubMed=1429593;
RA Bausch-Jurken M.T., Mahnke-Zizelman D.K., Morisaki T., Sabina R.L.;
RT "Molecular cloning of AMP deaminase isoform L. Sequence and bacterial
expression of human AMPD2 cDNA.";
RL J. Biol. Chem. 267:22407-22413(1992).
RN [2]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=96103174; PubMed=8526848;
RA Van den Bergh F., Sabina R.L.;
RT "Characterization of human AMP deaminase 2 (AMPD2) gene expression
reveals alternative transcripts encoding variable N-terminal
extensions of isoform L.";
RL Biochem. J. 312:401-410(1995).
RN [3]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=96350465; PubMed=8764830;
RA Mahnke-Zizelman D.K., van den Bergh F., Bausch-Jurken M.T., Eddy R.,
Sait S., Shows T.B., Sabina R.L.;
RT "Cloning, sequence and characterization of the human AMPD2 gene:
evidence for transcriptional regulation by two closely spaced
promoters.";
RL Biochim. Biophys. Acta 1308:122-132(1996).
CC -!- FUNCTION: AMP DEAMINASE PLAYS A CRITICAL ROLE IN ENERGY
METABOLISM.
CC -!- CATALYTIC ACTIVITY: AMP + H(2)O = IMP + NH(3).
CC -!- PATHWAY: PURINE NUCLEOTIDE CYCLE.
CC -!- SUBUNIT: Homotetramer.
CC -!- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=4;
Name=Ex1B-2-3;
Isold=Q01433-1; Sequence=Displayed;
Name=Ex1A-2-3;
Isold=Q01433-2; Sequence=VSP_001271, VSP_001272;
Name=Ex1A-3;
Isold=Q01433-3; Sequence=VSP_001274;
Name=Ex1B-3;
Isold=Q01433-4; Sequence=VSP_001273;
CC -!- TISSUE SPECIFICITY: THREE ISOFORMS ARE PRESENT IN MAMMALS: AMP
DEAMINASE 1 IS THE PREDOMINANT FORM IN SKELETAL MUSCLE; AMP

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M30449; AAA34420.1; -;
 DR EMBL; Z46659; CAA86620.1; -;
 DR PIR; S49744; S49744.
 DR SGD; S0004498; AMD1.
 DR GO; GO:0003876; P:AMP deaminase activity; IDA.
 DR GO; GO:006163; P:purine nucleotide metabolism; IC.
 DR InterPro; IPR006650; A/AMP_deam_sub.
 DR InterPro; IPR001365; A/AMP_deaminase.
 DR InterPro; IPR006329; AMP_deaminase.
 DR Pfam; PF00962; A_deaminase; 1.
 DR TIGRFAMs; TIGR01429; AMP_deaminase; 1.
 DR PROSITE; PS00485; A_DEAMINASE; 1.
 KW Hydrolase; Nucleotide metabolism.
 FT ACT_SITE 422 422 POTENTIAL.
 FT ACT_SITE 631 631 POTENTIAL.
 FT ACT_SITE 707 707 POTENTIAL.
 FT ACT_SITE 708 708 POTENTIAL.
 FT CONFLICT 568 568 F->C (IN REF. 1).
 SQ SEQUENCE 810 AA; 93301 MW; 746DCB43B9B45C93 CRC64;
 Query Match 53.1%; Score 1647; DB 1; Length 810;
 Best Local Similarity 53.7%; Pred. No. 2.9e-110;
 Matches 316; Conservative 83; Mismatches 152; Indels 38; Gaps 3;
 QY 12 VINDPCTPKPNPFTYVPEPKSEHVQF-----VDG----- 43
 DB 211 VQNDQNNKPKGWVYPPPKPSYNSDTKTVPVTKNPDAEVFOFTKCEIPGDPWDF 270
 QY 44 -----VIHYVADKCTESYVPVADATFTFDLHYILRVTAAGNTRVCHNRLNLEH 95
 DB 271 TLNDDSYVVRSGKTD--ELIAQPTLDYDLEKMSISSDGPFAKYRRLQYLEA 328
 QY 96 KFKFLMLNADREFLAQKTAHRDFYVNRKVDTHVHSACMQKHLRIKSKLRKEPDE 155
 DB 329 RNLYLLANEYQETSYSKENPHDFYVNRKVDTHVHSACMQKHLRIKSKLRKSKDE 388
 QY 156 VIFRDGTWTLKEVPESLDLTGYDLNVDLLDHADKSTFHRFDKPNLKNPCGQSLRE 215
 DB 389 KVIKFDGKLLTLDVFRSLHATGYDLSIDTLDGHAHKDTFHRFDKPNLKNPCIGESRLRE 448
 QY 216 IFKQDNLIQGRFLAELTKQVFSDLASKYQMAEYRISYGRKQSEWDOLASWVNNELH 275
 DB 449 IFKTNNTYIKGYLADITKQVFDLNSKYQCEYRISYGRSLDEWDKLSASWVINDKYI 508
 QY 276 SGNVVVLVQIPRLVNYKEMGIVTSFQNLDDNI FVPLFEVITDPAHQHLPOLHVLKQVGL 335
 DB 509 SHNRVLVQIPRLYDIYKYGIVQSFQDICKNLQFLEFVTKNPQSHKLVFLQVLTGF 568
 QY 336 DLVDESKEPERTHMPTPEQWTVNFPASYIYCYANLFTLNKLRSGKMTTIKFR 395
 DB 569 DSVDDSEKVDKDFHRKYPKPLWEAPQNPYPYLYLYNSVASLNQWRAKRGFTLVLR 628
 QY 396 PHAGEAGVDHLAAATFLCHNTSHGINLRKSPVLQVLYLQIGLAMSPLSNNSFLDYH 455
 DB 629 PHCAGEAGDPEHLVSAVLLAHGSHGILLRKKVPFQVLYLDQVGLAMSPSLNNAFLTYD 688
 QY 456 RNPFTFFRGNLVSLSTDPPDQIHLTKPELVEEYSIAASLWKLSCDCLCEARNVYQS 515
 DB 689 KNPPFRYFKGNLVLSTDDPLQFSYTRPLEEYSVAQIYKLNVDNCELARNVQLQS 748
 QY 516 GFSHALKAHWIKGVFKRGPAGNDIHRNTVPHIRVQFREMIRNEMKLV 564
 DB 749 GWEAQIKGHWIKGDFKSGVEGNDVVRVTVNPDIRINRYDYTLSTLELV 797
 RESULT 3

AMDM_SCHPO
 ID AMDM SCHPO STANDARD; PRT; 846 AA.
 AC P50998; Q9URV7;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 43, Last annotation update)
 DE AMP deaminase (EC 3.5.4.6) (Myoadenylate deaminase).
 GN ADAL OR SPBC106.04.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Rochet M., Levesque H., Gaillardin C.;
 RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
 RW [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Voiclaert G., Aert R., Robben J., Grymonprez B.,
 RA Weidjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Avee S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe."
 RL Nature 415:871-880(2002).
 CC -1- FUNCTION: AMP DEAMINASE PLAYS A CRITICAL ROLE IN ENERGY
 CC METABOLISM.
 CC -1- CATALYTIC ACTIVITY: AMP + H(2)O = IMP + NH(3).
 CC -1- PATHWAY: PURINE NUCLEOTIDE CYCLE.
 CC -1- SUBUNIT: Homotrimer (By similarity).
 CC -1- SIMILARITY: Belongs to the adenosine and AMP deaminases family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X91498; CAA62797.1; ALT_SEQ.
 DR EMBL; AL110295; CAB53720.1; -;
 DR PIR; T39261; T39261.
 DR GeneDB SPombe; SPBC106.04; -;
 DR InterPro; IPR006650; A/AMP_deam_sub.
 DR InterPro; IPR001365; A/AMP_deaminase.
 DR InterPro; IPR006329; AMP_deaminase.
 DR Pfam; PF00962; A_deaminase; 1.
 DR TIGRFAMs; TIGR01429; AMP_deaminase; 1.
 DR PROSITE; PS00485; A_DEAMINASE; 1.

```
KW Hydrolase; Nucleotide metabolism.
FT ACT SITE 394 394 POTENTIAL.
FT ACT SITE 603 603 POTENTIAL.
FT ACT SITE 679 679 POTENTIAL.
FT ACT SITE 680 680 POTENTIAL.
FT CONFLICT 483 520
FT
FT
SQ SEQUENCE 846 AA; 97454 MW; 701CC540A8E9C25 CRC64;

Query Match 52.2%; Score 1617; DB 1; Length 846;
Best Local Similarity 57.8%; Pred. No. 4.3e-108;
Matches 304; Conservative 84; Mismatches 126; Indels 12; Gaps 3;

QY 43 GVTHVYADKDC-----TESYIPVADATFTFDLHYILRVTAAGNTRVCHNRLNLEHKEPK 98
DB 245 GIPQVYENDSAYTAGTSPFH-IPTIRDYIDLEFLSASSDPSKSFRRLOYLEGRWN 303
QY 99 FHLMLNADREFLAQKTAHRDFYVNRKVDTHVHSSACMNQKHLRLRIKSLRKEPDEVVI 158
DB 304 MYMLNEXQELADTKKVPKHRDFYVNRKVDTHVHSSALANQKHLRLRIKSLRKEPDEVVI 363
QY 159 FRDGTWTKVEPESLDLTGYDLNVDLLDVLHADKSTFHRDPFNKYNPCGSRRLREIPL 218
DB 364 WRDGGKLTQEVFDSLSKLTSDYLSIDTLDHAAHTDTHFRDFKFNKYNPIGESRLRTIEL 423
QY 219 KODNLQGRFLAELTKQVFSLSASKYQMAEYRISYGRKQSEWDOLASWVNNELHSGN 278
DB 424 KTDNDINGRYLAELTKVEFDTLQTKYQMAEYRISYGRNREWDKLAAWIDNELFSN 483
QY 279 VVWLQIPRLYNYKEMGVTSFQNLNDNI FVPLFEVITDPASHPOLHVFQLKQVGLDLV 338
DB 484 VRWLQIPRLYNYKSGIVETEEVNRVNFELFEVTKDPRTHPKLHVLQVIGFDSV 543
QY 339 DDESKPERPTKMPPEQNTVNFNPAFYAYCYANLFTLNKLRSGKMTTKFRPHR 398
DB 544 DDESKPERPTKMPPEQNTVNFNPAFYAYCYANLFTLNKLRSGKMTTKFRPHR 603
QY 399 GEAGDVHDLAATLLCHNLSHGINKRSPVLYLYLGOIGQAMSPLSNLSFLDVRNP 458
DB 604 GEAGDVHDLAATLLCHNLSHGINKRSPVLYLYLGOIGQAMSPLSNLSFLDVRNP 663
QY 459 FPTFFORGLNVLSTDDPLQIHILTKPELVEEYSIAASLKLSCDLCIARNSVYOSGFS 518
DB 664 FLTYFKRGLNVLSTDDPLQFAFTREPLIEEYAVAAQIKLSAVDMCELARNSVLQSGFE 723
QY 519 HALKAHWIGKNYKFGPAGNDIHRTPVHVRVQFREMWRNEMKLV 564
DB 724 RQLKERWLVDF-----QDIDRTNVPIRLAYRALTLTQETALV 762

RESULT 4
ID _HUMAN3 HUMAN STANDARD; PRT; 767 AA.
AC Q01432;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE AMP deaminase 3 (EC 3.5.4.6) (AMP deaminase isoform E) (Erythrocyte
DE AMP deaminase).
GN AMPD3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93042002; PubMed=1420359;
RA Yamada Y., Goto H., Ogasawara N.;
RT "Cloning and nucleotide sequence of the cDNA encoding human
RL erythrocyte-specific AMP deaminase.";
RN [2]
Biochim. Biophys. Acta 1171:125-128 (1992).
```

```
RP SEQUENCE FROM N.A.
RX TISSUE=Keratinocytes;
RX MEDLINE=93015995; PubMed=1400401;
RA Mahnke-Zielman D.K., Sabina R.L.;
RT "Cloning of human AMP deaminase isoform E cDNAs. Evidence for a third
RT AMP gene exhibiting alternatively spliced 5'-exons.";
RL J. Biol. Chem. 267:20866-20877 (1992).
RN [3]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=96201708; PubMed=8611627;
RA Mahnke-Zielman D.K., Eddy R., Shows T.B., Sabina R.L.;
RT "Characterization of the human AMPD3 gene reveals that 5' exon usage
RT is subject to transcriptional control by three tandem promoters and
RT alternative splicing.";
RL Biochim. Biophys. Acta 1306:75-92 (1996).
RN [4]
RP SEQUENCE FROM N.A., AND VARIANT CVS-573.
RX MEDLINE=94272475; PubMed=8004104;
RA Yamada Y., Goto H., Ogasawara N.;
RT "A point mutation responsible for human erythrocyte AMP deaminase
RT deficiency.";
RL Hum. Mol. Genet. 3:331-334 (1994).
CC -!- FUNCTION: AMP DEAMINASE PLAYS A CRITICAL ROLE IN ENERGY
CC METABOLISM.
CC -!- CATALYTIC ACTIVITY: AMP + H(2)O = IMP + NH(3) .
CC -!- PATHWAY: PURINE NUCLEOTIDE CYCLE.
CC -!- SUBUNIT: Homotetramer.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1B;
CC IsoId=Q01432-1; Sequence=Displayed;
CC Name=1A;
CC IsoId=Q01432-2; Sequence=VSP_001275, VSP_001277;
CC Name=1C;
CC IsoId=Q01432-3; Sequence=VSP_001276, VSP_001278;
CC -!- TISSUE SPECIFICITY: THREE ISOFORMS ARE PRESENT IN MAMMALS: AMP
CC DEAMINASE 1 IS THE DOMINANT FORM IN SKELETAL MUSCLE; AMP
CC DEAMINASE 2 PREDOMINATES IN SMOOTH MUSCLE, NON-MUSCLE TISSUE,
CC EMBRYONIC MUSCLE AND UNDIFFERENTIATED MYOBLASTS; AMP DEAMINASE 3
CC IS FOUND IN ERYTHROCYTES.
CC -!- SIMILARITY: Belongs to the adenosine and AMP deaminases family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC -----
CC EMBL; D12775; BAA02240.1; -
CC EMBL; M84720; AAS58365.1; -
CC EMBL; M84721; AAS58366.1; -
CC EMBL; M84722; AAS58367.1; -
CC EMBL; U29926; AAB60410.1; -
CC EMBL; U29929; AAB60410.1; JOINED.
CC EMBL; U29907; AAB60410.1; JOINED.
CC EMBL; U29909; AAB60410.1; JOINED.
CC EMBL; U29910; AAB60410.1; JOINED.
CC EMBL; U29911; AAB60410.1; JOINED.
CC EMBL; U29916; AAB60410.1; JOINED.
CC EMBL; U29917; AAB60410.1; JOINED.
CC EMBL; U29918; AAB60410.1; JOINED.
CC EMBL; U29922; AAB60410.1; JOINED.
CC EMBL; U29924; AAB60410.1; JOINED.
CC EMBL; U29925; AAB60410.1; JOINED.
CC EMBL; U29926; AAB60408.1; -
CC EMBL; U29912; AAB60408.1; JOINED.
CC EMBL; U29929; AAB60408.1; JOINED.
CC EMBL; U29907; AAB60408.1; JOINED.
CC EMBL; U29909; AAB60408.1; JOINED.
CC EMBL; U29910; AAB60408.1; JOINED.
CC EMBL; U29911; AAB60408.1; JOINED.
CC EMBL; U29916; AAB60408.1; JOINED.
```


This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: U90888; AAC53348.1; ALT INIT.
InterPro: IPR006650; A/AMP deam. sub.
InterPro: IPR001365; A/AMP deaminase.
InterPro: IPR006329; AMP deaminase.
Pfam: PF00962; A deaminase; 1.
TIGRFAMs: TIGR01429; AMP deaminase; 1.
PROSITE: PS00485; A DEAMINASE; 1.
KW Hydrolase; Nucleotide metabolism; Multigene family;
KW Alternative splicing.
SQ SEQUENCE 765 AA; 88483 MW; F826FA5966221B09 CRC64;
Query Match 49.5%; Score 1534; DB 1; Length 765;
Best Local Similarity 51.8%; Pred. No. 3.3e-102;
Matches 288; Conservative 90; Mismatches 164; Indels 14; Gaps 4;
QY 19 PKPNPFTYVPEPKS-EHVFQTVGVHVIYADKDCTE-----SIYFVADATFTFDLHYI 73
Db 200 PLPQEDPYCLDDAPPNLGVLVRMGQVLFVYDQNTMLERQEPHSLPYDPDLTKYIVDMSHI 259
QY 74 LRVTAAGNTRTYCHNRLLNLEHKFKFHLMLNADREFLAQKTAHRDFYNNVRKVDTHVHS 133
Db 260 LALITDGPCTKYCHRRLLNLEHKFKFHLMLNADREFLAQKTAHRDFYNNVRKVDTHVHS 319
QY 134 ACNQKHLLRFKSKLRKPEDEVVIFRDGTMTLKVFESLDITGVLDVLLDVHADKS 193
Db 320 ACNQKHLLRFKSKLRKPEDEVVIFRDGTMTLKVFESLDITGVLDVLLDVHADKS 379
QY 194 TPRPDKFNLYKYNPCGSLREIFLQDNLIQGRFLAELTKQVFDLSASKYQMAEYRIS 253
Db 380 TPGHDFKFNLYKYNPCGSLREIFLQDNLIQGRFLAELTKQVFDLSASKYQMAEYRIS 439
QY 254 IYGRKSEWDLASWVNNELHSGNVVWLVQIPRLNIVYKEMGIVTSFQNLNIIIVPLF 313
Db 440 IYGRSPKWSLARWTFQHKYVFNMRWIIQVPRYIDFRSKKLLPSFGMLNIFLPLF 499
QY 314 EVTIDPASHPOLHVFLLKVVGLDLDVDESKPERRPTKHM-----PTPEQWTVNFNPAFSY 368
Db 500 QATINQDREHLLFLKYVTFGSVDDESK-----HSDHMFSDKSPFDLWTSEQNPPYSY 555
QY 369 YAYCYVANFLTKLRESKGMITTKPRPHAGEAGVDVHDLAATFLCHNISHGILNRKSPV 428
Db 556 YLYVMYANIMVNLNRRERGLSTFLFRPHCGEAGSITHLVSAFLTADNISHGLLLKKSPV 615
QY 429 LQVLYLQGLGLAMSPLSNNLSFLDYHRNPDPFTFFQFQRLNVLSTDDPLQIHLTKELPVE 488
Db 616 LQVLYLQGLGLAMSPLSNNLSFLDYHRNPDPFTFFQFQRLNVLSTDDPLQIHLTKELPVE 675
QY 489 EYSLASLWKLSSCDLCEIARNVYSGSGFHALKAHWIGKYNPKRGPAGNDIHRVTNVPHI 548
Db 676 EYSLASLWKLSSCDLCEIARNVYSGSGFHALKAHWIGKYNPKRGPAGNDIHRVTNVPHI 735
QY 549 RVQFREMIRWENMKLV 564
Db 736 RMAFRVETLCLNELSFL 751

RESULT 6
AMD3 MOUSE

ID AMD3 MOUSE STANDARD; PRT; 766 AA.
AC O08739; O88692;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-NOV-2000 (Rel. 39, Last annotation update)
DE AMP deaminase 3 (EC 3.5.4.6) (AMP deaminase isoform E) (AMP deaminase
H-type) (Heart-type AMPD).

GN AMPD3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=97279027; PubMed=9133604;
RA Wang X., Morisaki H., Sermuivitayawong K., Mineo I., Toyama K.,
Ogaawara N., Mukai T., Morisaki T.;
RT "Cloning and expression of cDNA encoding heart-type isoform of AMP
deaminase.";
RL Gene 188:285-290 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RX MEDLINE=97468276; PubMed=9321472;
RA Sermuivitayawong K., Wang X., Nagabukuro A., Matsuda Y., Morisaki H.,
Toyama K., Mukai T., Morisaki T.;
RT "Genomic organization of Ampd3, heart-type AMPD gene, located in
mouse chromosome 7.";
RL Mamm. Genome 8:767-769 (1997).
CC -!- FUNCTION: AMP DEAMINASE PLAYS A CRITICAL ROLE IN ENERGY
METABOLISM.
CC -!- CATALYTIC ACTIVITY: AMP + H(2)O = IMP + NH(3).
CC -!- PATHWAY: PURINE NUCLEOTIDE CYCLE.
CC -!- SUBUNIT: Homotetramer (By similarity).
CC -!- TISSUE SPECIFICITY: FOUND IN HEART, LUNG BRAIN, SPLEEN, KIDNEY AND
TO A LESSER EXTENT IN LIVER.
CC -!- SIMILARITY: Belongs to the adenosine and AMP deaminases family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL: D85596; BAA19933.1; ALT SEQ.
EMBL: D88994; BAA32548.1; JOINED.
EMBL: D88985; BAA32548.1; JOINED.
EMBL: D88986; BAA32548.1; JOINED.
EMBL: D88987; BAA32548.1; JOINED.
EMBL: D88988; BAA32548.1; JOINED.
EMBL: D88989; BAA32548.1; JOINED.
EMBL: D88990; BAA32548.1; JOINED.
EMBL: D88991; BAA32548.1; JOINED.
EMBL: D88992; BAA32548.1; JOINED.
EMBL: D88993; BAA32548.1; JOINED.
MGD; MGI:1096344; Ampd3.
InterPro: IPR006650; A/AMP deam. sub.
InterPro: IPR001365; A/AMP deaminase.
InterPro: IPR006329; AMP deaminase.
Pfam: PF00962; A deaminase; 1.
TIGRFAMs: TIGR01429; AMP deaminase; 1.
PROSITE: PS00485; A DEAMINASE; 1.
KW Hydrolase; Nucleotide metabolism; Multigene family.
KW CONFLICT 251 251 K -> E (IN REF. 2).
SQ SEQUENCE 766 AA; 88665 MW; 949B613565EFAAA0 CRC64;

Query Match 49.4%; Score 1531; DB 1; Length 766;
Best Local Similarity 51.6%; Pred. No. 5.4e-102;
Matches 287; Conservative 91; Mismatches 164; Indels 14; Gaps 4;
QY 19 PKPNPFTYVPEPKS-EHVFQTVGVHVIYADKDCTE-----SIYFVADATFTFDLHYI 73
Db 201 PLPQEDPYCLDDAPPNLGVLVRMGQVLFVYDQNTMLERQEPHSLPYDPDLTKYIVDMSHI 260
QY 74 LRVTAAGNTRTYCHNRLLNLEHKFKFHLMLNADREFLAQKTAHRDFYNNVRKVDTHVHS 133
Db 261 LALITDGPCTKYCHRRLLNLEHKFKFHLMLNADREFLAQKTAHRDFYNNVRKVDTHVHS 320

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCB_TaxID=9606;
 RP [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90264442; PubMed=23451176;
 RA Sabina R.L., Morisaki T., Clarke P., Eddy R., Shows T.B., Morton C.C.,
 RA Holmes E.W.;
 RT "Characterization of the human and rat myoadenylate deaminase
 RT genes.";
 RL J. Biol. Chem. 265:9423-9433 (1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92131279; PubMed=1370861;
 RA Sabina R.L., Fishbein W.N., Pezeshkpour G., Clarke P.R., Holmes E.W.;
 RT "Molecular analysis of the myoadenylate deaminase deficiencies.";
 RL Neurology 42:170-179 (1992).
 RN [3]
 RP VARIANT LEU-48.
 RX MEDLINE=92335317; PubMed=1631143;
 RA Morisaki T., Gross M., Morisaki H., Pongratz D., Zoellner N.,
 RA Holmes E.W.;
 RT "Molecular basis of AMP deaminase deficiency in skeletal muscle.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:6457-6461 (1992).
 RN [4]
 RP VARIANTS AMPD DEFICIENCY TRP-388 AND HIS-425.
 RX MEDLINE=20556569; PubMed=11102975;
 RA Morisaki H., Higuchi I., Abe M., Osame M., Morisaki T.;
 RT "First missense mutations (R388W and R425H) of AMPD1 accompanied with
 RT myopathy found in a Japanese patient.";
 RL Hum. Mutat. 16:467-472 (2000).
 CC -1- FUNCTION: AMP DEAMINASE PLAYS A CRITICAL ROLE IN ENERGY
 CC METABOLISM.
 CC -1- CATALYTIC ACTIVITY: AMP + H(2)O = IMP + NH(3).
 CC -1- PATHWAY: PURINE NUCLEOTIDE CYCLE.
 CC -1- SUBUNIT: Homotetramer.
 CC -1- TISSUE SPECIFICITY: THREE ISOFORMS ARE PRESENT IN MAMMALS: AMP
 CC DEAMINASE 1 IS THE PREDOMINANT FORM IN SKELETAL MUSCLE; AMP
 CC DEAMINASE 2 PREDOMINATES IN SMOOTH MUSCLE, NON-MUSCLE TISSUE,
 CC EMBRYONIC MUSCLE AND UNDIFFERENTIATED MYOBLASTS; AMP DEAMINASE 3
 CC IS FOUND IN ERYTHROCYTES.
 CC -1- DISEASE: Defects in AMPD1 are the cause of AMP deaminase
 CC deficiency (AMPD deficiency). It results in exercise-related
 CC myopathy characterized by muscle aches, cramps, and early fatigue.
 CC It is one of the most common inherited defects in the Caucasians,
 CC but not in Asians.
 CC -1- SIMILARITY: Belongs to the adenosine and AMP deaminases family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M37931; AAG24258.1; -
 DR EMBL; M37920; AAG24258.1; JOINED.
 DR EMBL; M37921; AAG24258.1; JOINED.
 DR EMBL; M37922; AAG24258.1; JOINED.
 DR EMBL; M37923; AAG24258.1; JOINED.
 DR EMBL; M37924; AAG24258.1; JOINED.
 DR EMBL; M37927; AAG24258.1; JOINED.
 DR EMBL; M37928; AAG24258.1; JOINED.
 DR EMBL; M37929; AAG24258.1; JOINED.
 DR EMBL; M37930; AAG24258.1; JOINED.
 DR EMBL; M60092; AAG24258.1; JOINED.
 DR PIR; I39444; I39444.
 DR Genew; HGNC:468; AMPD1.
 DR MIM; 102770; -
 DR GO; GO:0003876; F:AMP deaminase activity; TAS.
 DR GO; GO:0006197; P:adenylate deaminase reaction; TAS.

DR InterPro; IPR006650; A/AMP deam sub.
 DR InterPro; IPR001365; A/AMP deaminase.
 DR InterPro; IPR006329; AMP deaminase.
 DR Pfam; PF00962; A_deaminase; 1.
 DR TIGRfams; TIGR01429; AMP_deaminase; 1.
 DR PROSITE; PS00485; A_DEAMINASE; 1.
 KW Hydrolase; Nucleotide metabolism; Multigene family; Polymorphism;
 Disease mutation.
 FT ACT_SITE 363 BY SIMILARITY.
 FT ACT_SITE 573 BY SIMILARITY.
 FT ACT_SITE 649 BY SIMILARITY.
 FT ACT_SITE 650 BY SIMILARITY.
 FT VARIANT 48 P -> L (POLYMORPHISM; ACTIVITY COMPARABLE
 TO WILD TYPE).
 FT /FTID=VAR 013270.
 FT FT
 FT R -> W (IN AMPD DEFICIENCY; LOSS OF
 ACTIVITY).
 FT /FTID=VAR 013271.
 FT FT
 FT R -> H (IN AMPD DEFICIENCY; LOSS OF
 ACTIVITY).
 FT /FTID=VAR 013272.
 FT FT
 SQ SEQUENCE 747 AA; 86489 MW; 1E15EBE98B95763 CRC64;
 Query Match 48.3%; Score 1497; DB 1; Length 747;
 Best Local Similarity 51.4%; Pred. No. 1.4e-99;
 Matches 289; Conservative 89; Mismatches 162; Indels 22; Gaps 8;
 QY 18 TP--KPENPFP--TVPEPKSEHVQTVGVVHVY-----ADKDCETSI-YPVADATTF 67
 DB 185 TPPVKKGEDPRTDNLNENLGLYHL-KMKDGVVYVNEAAVSKDPKPLPVNLD--TFL 241
 QY 68 TDLHYILRTAAGNTRTVCHNRLNLEHKFKFHLMLNADREFLAQCTAPHRDFYVNRKVD 127
 DB 242 DMNFLLALIAQGPVKTTHRLKFLSKFQVHQLNEMDELKELKNPHRDFYVNRKVD 301
 QY 128 THVHSACNQKHLRLFKSLKRKEPDEVIFRDTGYMTLKEVPSLDTGYDLNVLLD 187
 DB 302 THIAAACMNQKHLRLFKSKYQIDADRVVYSTKEKNLTLELFAKLKHPYDLTVDSL 361
 QY 188 VHADKSTHTRDPKFNLYNPGOSBLRIFLKQDNLIQGRFLAELTKOVFSLDSASKYQM 247
 DB 362 VHAGQTFQRFKFNKYNPVGASLRLDLYLKTNYINGEYPATIIKEVGADLVKAYQH 421
 QY 248 AEYRISYIGRKQSEWDQLASIVNNELHSGNVNVLVQIPRLYVNYKEMGIVTSFQNL 307
 DB 422 AEPLSIYIGRSPDESKLSMFVCHNHCNPNWMIQVRIYDVFRSKNFLPHFGKMLEN 481
 QY 308 IFVPLFEVTIDPASHPOLHVLKQVVGDLVDDESKPERRPTKHM-----PTPEQWTVNF 362
 DB 482 IFMPVFEATINQADPELSVFLKHITGFDSDVDESK---HSGHMFSSKSPKQEWTLK 537
 QY 363 NPAFSYIAYCYANUFTLNKLESKGMTIKRPHAGAGVDHLLAATFLLCHNISHGIN 422
 DB 538 NPSYTYIAYYVYANIMVLNLSRKRGMTFLRPHCGEAGALTHLMTAFMIADDDSHGLN 597
 QY 423 LKSPVLQVLYLGOIGLAMSPLSNLSFLDYHRNPFPTFFORGLNVLSTDDPLQIHLT 482
 DB 598 LKSPVLQVLYLQVLAQPIAMSPLSNLSFLDYHRNPFPTFFORGLNVLSTDDPLQIHLT 657
 QY 483 KEPLVEEYSIAASLWKLSCDLCEIARNVSVOGFSHALKAHWIGNYFKRGPAGNDIHR 542
 DB 658 KEPLMEEYAIQAQVFKLSTCDMCEVARNVSLCCGISHEEKVKFLGDNVLEEGPAGNDIRR 717
 QY 543 TNVPHIRVQFREMIRNENKLV 564
 DB 718 TNVAQIRWAYRYETWCYELNLI 739
 RESULT 9
 YJHO_YEAST
 ID YJHO_YEAST PRT; 888 AA.
 AC P40361;
 DT 01-FEB-1995 (Rel. 31, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 104.3 kDa protein in SMC3-MRPL8 intergenic region.
 GN YJL070C OR J1095 OR HRD888.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c;
 RX MEDLINE=95282514; PubMed=7762302;
 RA Vandenbol M., Durand P., Dion C., Portetelle D., Hilger F.;
 RT "Sequence of a 17.1 kb DNA fragment from chromosome X of
 Saccharomyces cerevisiae includes the mitochondrial ribosomal protein
 L8.";
 RL Yeast 11:57-60(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c;
 RA Sor F.J.;
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: Belongs to the adenosine and AMP deaminases family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z34288; CAA84052.1; -
 DR EMBL; Z49345; CAA89362.1; -
 DR EMBL; Z49344; CAA89361.1; -
 DR EMBL; X88851; CAA61309.1; -
 DR PIR; S50801; S50801.
 DR SGD; S0003606; YJL070C.
 DR InterPro; IPR001365; A/AMP deaminase.
 DR Pfam; PF00962; A deaminase; 1.
 KW Hypothetical protein; Hydrolase.
 SQ SEQUENCE 888 AA; 104263 MW; DB11A8086224114D CRC64;

 Query Match 20.2%; Score 627; DB 1; Length 888;
 Best Local Similarity 30.1%; Pred. No. 3.2e-37;
 Matches 184; Conservative 94; Mismatches 207; Indels 126; Gaps 21;
 QY 66 FETDLHYILRVTAAGNTRVCHNRLNLEHKKFHLMLNADREFLAQTAHPDRDFYVRK 125
 DB 281 FRDDFAYIIELIQSHKFNVEVRKLSYLLDKPFLFQYLSKKEILANKVVPYRDFYNGRK 340
 QY 126 VTHVHSAACMOKHLLRIKSKLRKEPDEVVIFRD---GTMTLKEVF----- 171
 DB 341 VRDLSLSGCIQRQLSEYIWEKINLEPRIV-YQDPETSKLSLRDIFQFCSSNDQPI 399
 QY 172 -----ESLDLTGYDLNVDLADVADKSTFHRFDKFNKYNPCGQSRLEIFLKQDN 222
 DB 400 AIGLKLIDDEFLD---WYRNIVLIDVHLTPNKVAKLVGKEMF-----YLLAKVLEFPD 451
 QY 223 LIQGRFLAEL-TQVPSDLSASKYQAEYRIS--TYGRKQSEWDQLASHVINNEHLSGNV 279
 DB 452 FIEGELYLAIFIKYVIHILEKSKYQLAQVSNVFPQYSSGEDYKFKFSQWLLRWKLVSNI 511
 QY 280 VMLVQIPRLY-NVYKEMGVITVFNQLLDNIFVPLF-----EVTIDPASHPQ---LHVFLK 330
 DB 512 RNNIQUIARIFPKLFE-NVVSFQFELDLINPLFTLEKEQIPDISSVNTDIIGQFLLS 570
 QY 331 QVVGDLVDDESKEP--RRPTKMTPEQWT-NVFNPAFSYYAYCYANLFTLNKLRSEK 387
 DB 571 NVCSMDLVKESDEYVWKEFTDMNCKPKFTWAGDNPTVAHYWYIYKSLAKVNFRLSQN 630
 QY 388 GMTTIKFRPHAGE-----AGDVHDLAATFLLCHNISHGINLRKSP----- 427

DB 631 LQNTITLRYCSPSSRTSQFGVDLYFTDQVESLVNCLLCC-----NGGLQVEPLWDAT 686
 QY 428 VLQYLYLQIGQIGLAMSPLSNLSL-----FL----- 452
 DB 687 MIQYLYFYLFQIPILAAPLSSVLSLNSQKSTFLKNKNVLEHDYLDKQETAKINPSRDITV 746
 QY 453 ---DYHRNPPTFPFQRLNVLSLTDPL-QIHLLKEPLVEEYSIAASLWKLSSCDLCEI 507
 DB 747 GEQSYETNPFPKMPFQMGKISLSKSLVNSYTTLEPLIEBYSVAASIIYLLNPTDLCEL 806
 QY 508 ARNSVYQSGFSHALKAHWIGNK-----YKRGPGAGND-----IHRTNVPHIRVOPR 553
 DB 807 SRTSVLSGSGYGVYKAHWIGVGVKAPYFENVGGIDWYDTAKDTSIKHNVPIMRRYR 866
 QY 554 EMIMRNEMKLV 564
 DB 867 KETLDQEWNV 877

 RESULT 10
 YB9Z YEAST
 ID YB9Z YEAST STANDARD; PRT; 797 AA.
 AC P38150;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DE Hypothetical 92.9 kDa protein in SSH1-APE3 intergenic region.
 GN YBR284W OR YBR2021.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c;
 RX MEDLINE=94378722; PubMed=8091861;
 RA Holmstroem K., Brandt T., Kalliesoe T.;
 RT "The sequence of a 32,420 bp segment located on the right arm of
 chromosome II from Saccharomyces cerevisiae.";
 RL Yeast 10:947-962(1994).
 CC -!- SIMILARITY: Belongs to the adenosine and AMP deaminases family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X76053; CAA53647.1; -
 DR EMBL; Z36153; CAA85248.1; -
 DR PIR; S44546; S44546.
 DR SGD; S0000488; YBR284W.
 DR InterPro; IPR001365; A/AMP deaminase.
 DR Pfam; PF00962; A deaminase; 1.
 KW Hypothetical protein; Transmembrane; Hydrolase.
 FT TRANSMEM 627 647 POTENTIAL
 SQ SEQUENCE 797 AA; 92903 MW; BF854E93D0B2C747 CRC64;

 Query Match 20.2%; Score 625; DB 1; Length 797;
 Best Local Similarity 30.6%; Pred. No. 3.9e-37;
 Matches 182; Conservative 86; Mismatches 213; Indels 114; Gaps 21;
 QY 66 FETDLHYILRVTAAGNTRVCHNRLNLEHKKFHLMLNADREFLAQTAHPDRDFYVRK 125
 DB 219 FREDFEWCLKTIRDSLSRFSSEKQLVNVKFPVQHLHSKEEMRQSKVPHKDFYNGRK 278
 QY 126 VTHVHSAACMOKHLLRIKSKLRKEPDEVV--IFRDGYMTLKEVF-SLDLTG----- 178
 DB 279 IDLMLLSGCFSQQLTFEFTWKURKEPDRVIHQAF-NGSHITLSQLFKVNFETGQFFN 337
 QY 179 -----YDLNVLDLADVADKSTFHRFDKFNKYNPCGQSRLEIFL 218

Db 338 GLKIIDDSFLEWVKVIYLAHYLVNDEMBIHTGS-----HGKQLRY-----YLIAKTEL 386
QY 219 KODNLIQGRFLAELTQK-VPSDLASQYQMAEYRISI-----YGRQOSE--WDQLASWTV 270
Db 387 EFDNYINGBYLAELLTFLIKPQESKYQCLQSLVDFQYLYHYDSDVDNWMVFWANWLN 446
QY 271 NNEHSGNVWLVQVIRPLYNVYKMGIVTSFQNLNDNI FVPLF--EVTIDPASHPOLHVF 328
Db 447 HYNIFNNIRWIRISRIYPELYHTKVKNFQYELNLIKPLFNAENYHLKSLPILKPF 506
QY 329 LKQWGLDLV--DOESKPERRTKHMPTPEQWTN-VFNPAFSYAYCYANFLTNKLRE 385
Db 507 LSQVSSIDLQIDTDNVIWKNFTAVSCLPKDMTSGGDNFTISQYMYVYVNLTKLNIHQ 566
QY 386 -----SKGWTIIFRPHAGEAGDVHDLAATFL-CHNISHGINL 423
Db 567 ALHONTFTLRSSCSPTSMRTSQFSNTLNFTF-----TEAILNNFLLACGFLNAEVL 620
QY 424 -RKSPVLQYLYLGOIGLAMSPLSN-----NSLFLD----YHRNPPPTFFQ 464
Db 621 WNAFSLVLYLFSQIPMVVAPLNSIVDSKPTMLQCAQTGLVLEPSKYKKNPFMKPFE 680
QY 465 RGNVSLSTDDPLQIH-LTKEPLVEEYSIAASLWKLSSCDLCEIARNVSYQSGFHALKA 523
Db 681 MGFKISLSSESILYNNSTKEPIIEYSVAASIYRLHSADLCCELLNSVITSFGSSTLKN 740
QY 524 HWIG-----KNVFKRGA-----GNDIHTNVPHIRVQPREMIWENMKLV 564
Db 741 KMLGVSLASHDYEVNTGFGWKYDCKPNTSLHNVPIIRQYRSSTLAGEWRLI 795

RESULT 11

AMD2_RAT ID AMD2_RAT STANDARD; PRT; 88 AA.
AC Q02356;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE AMP deaminase 2 (EC 3.5.4.6) (AMP deaminase isoform L) (Fragment).
GN AMPD2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=90307656; PubMed=2365682;
RA Morieaki T., Sabina R.L., Holmes E.W.;
RT "Adenylylate deaminase. A multigene family in humans and rats.";
RL J. Biol. Chem. 265:11482-11486(1990).
CC -!- FUNCTION: AMP DEAMINASE PLAYS A CRITICAL ROLE IN ENERGY METABOLISM.
CC -!- CATALYTIC ACTIVITY: AMP + H(2)O = IMP + NH(3).
CC -!- PATHWAY: PURINE NUCLEOTIDE CYCLE.
CC -!- SUBUNIT: Homotetramer.
CC -!- TISSUE SPECIFICITY: THREE ISOFORMS ARE PRESENT IN MAMMALS: AMP DEAMINASE 1 IS THE PREDOMINANT FORM IN SKELETAL MUSCLE; AMP DEAMINASE 2 PREDOMINATES IN SMOOTH MUSCLE, NON-MUSCLE TISSUE, EMBRYONIC MUSCLE AND UNDIFFERENTIATED MYOBLASTS; AMP DEAMINASE 3 IS FOUND IN ERYTHROCYTES.
CC -!- SIMILARITY: Belongs to the adenosine and AMP deaminases family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M38126; AAA40728.1; -.

DR PIR; A37056; A37056.
DR InterPro; IPR006650; A/AMP deam sub.
DR InterPro; IPR001365; A/AMP deaminase.
DR Pfam; PF00962; A_deaminase; 1.
DR PROSITE; PS00485; A_DEAMINASE; 1.
KW Hydrolyase; Nucleotide metabolism; Multigene family.
FT NON_TER 1
FT ACT_SITE 3 POTENTIAL.
FT ACT_SITE 79 POTENTIAL.
FT ACT_SITE 80 POTENTIAL.
FT NON_TER 88
SQ SEQUENCE 88 AA; 9863 MW; 1AF8DC091DB298E7 CRC64;
Query Match 11.3%; Score 350; DB 1; Length 88;
Best Local Similarity 72.7%; Pred No. 1e-18;
Matches 64; Conservative 7; Mismatches 17; Indels 0; Gaps 0;
QY 396 PHAGSAGVDVHDLAATFLCHNISHGINIRKSPVLYLYLGOIGLAMSPLSNNSFLDYH 455
Db 1 PHCGSAGPIHLLVSFAFLAENISHGLLRKAPVLYLYLAQIGIAMSPLSNNSFLSYH 60
QY 456 RNPPTFFQRLNVLSTDDPLQIHLYTK 483
Db 61 RNPPEYLSRGLMVSLSLTDPLQPHFTK 88
RESULT 12
ADD_SHEON ID ADD_SHEON STANDARD; PRT; 331 AA.
AC Q8EBD4;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Adenosine deaminase (EC 3.5.4.4) (Adenosine aminohydrolase).
GN ADD OR SO4731.
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MR-1;
RX MEDLINE=2297686; PubMed=12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
RA Mueller J., Kouri H., Gill J., Uterback T.R., McDonald L.A., C.M.;
RA Feldblyum T.V., Smith H.O., Venter J.C., Nealsen K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium Shewanella oneidensis.";
RL Nat. Biotechnol. 20:1118-1123(2002).
CC -!- CATALYTIC ACTIVITY: Adenosine + H(2)O = inosine + NH(3).
CC -!- SIMILARITY: Belongs to the adenosine and AMP deaminases family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF015905; RAN57690.1; -.
DR TIGR; SO4731; -.
DR HAMAP; MF 00540; -; 1.
DR Pfam; PF00962; A_deaminase; 1.
DR TIGRFAMs; TIGR01430; aden_deam; 1.
DR PROSITE; PS00485; A_DEAMINASE; FALSE NEG.
KW Hydrolyase; Nucleotide metabolism; Complete proteome.
FT ACT_SITE 197 POTENTIAL.

FT ACT SITE 245 245 POTENTIAL.
 FT ACT SITE 278 278 POTENTIAL.
 FT ACT SITE 279 279 POTENTIAL.
 SQ SEQUENCE 331 AA; 36175 MW; 5304F503789C73CF CRC64;

Query Match 4.8%; Score 150; DB 1; Length 331;
 Best Local Similarity 22.8%; Pred. No. 0.0012;
 Matches 90; Conservative 49; Mismatches 122; Indels 134; Gaps 19;

QY 182 NVDLVDVHADKSTFRFDNFNLKYNPCGO-SRLREIFLKQDNL-----IQGR----- 227
 DB 5 SIPLVDL-----HHLDGNVRVNTIWEHGCHGIALPADSLETLPVQIQKETSIV 57
 QY 228 -FLAEL--TKQVFSLSAK-----YMAEYRISYGRKQSEWDQLASIV 270
 DB 58 AFLKLDVMAVLADLVKRIAYENVADAALSGLDYAEIRFSY-----YMA 106

QY 271 NNEHSGNVVLMVQIPRLYNYVKEMGIVTSFQNLNDNFVPL--FEVTD----- 318
 DB 107 NKKLPTEGV-----BAVIDGVKAGLDKYQVNNILGIMRSFG 145

QY 319 -PASHPQLHVFL---KQVYGLDLDVDESKPERRPTKHMPTPEQWTVNFVPAFSYAYCY 374
 DB 146 QPACTQELSGLLAHKQHLVAMDLAGDE-----LGFGGE---LFNEHF----- 184

QY 375 ANLFTLKURESGKMTTIKFRPHAGEAD-----VDHLAATFLCHNISHGINLRKSP 427
 DB 185 -----KVRDA---GLAITAHAGEAGSQQSMQAIQELGAT-----RIGHGVNAIHP 229

QY 428 VLQYLYLQIGLAMSPLSN--NSLFLDYHRPPFPFFORGLNVSLSTDPPQIHLTKRP 485
 DB 230 KLMVEYLAKRHIGIESCPTSNLHTSTVSSYAEHFFRTFMDAGVLGLNTDDP---GVSAID 286

QY 486 LVEEYSIAASLWKLSCDCEIARNVYQSGFSHA 520
 DB 287 INKHYRIAKFELGSLDAELAQVRGVEAFUSE 321

RESULT 13
 ADD1_STRCO STANDARD; PRT; 334 AA.

AC Q97EV1.
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Adenosine deaminase (EC 3.5.4.4) (Adenosine aminohydrolase).
 GN ADD OR CAC3005.
 OS Clostridium acetobutylicum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1488;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
 RX MEDLINE=21359325; PubMed=11466286;
 RA Noelling J., Breton G., Omeichenko M.V., Makarova K.S., Zeng Q.,
 RA Gibson R.L., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
 RA Tatunov R.L., Sabathe F.J., Doucette-Stamm L., Soucaille P., Daly M.J.,
 RA Bennett G.N., Koonin E.V., Smith D.R.;
 RA "Genome sequence and comparative analysis of the solvent-producing
 RT bacterium Clostridium acetobutylicum."
 RL J. Bacteriol. 183:4823-4838(2001).
 CC -!- CATALYTIC ACTIVITY: Adenosine + H(2)O = inosine + NH(3).
 CC -!- SIMILARITY: Belongs to the adenosine and AMP deaminases family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AE007797; AAK80946.1; -.
 DR PIR; G97269; G97269.
 DR HAMAP; MF_00540; -; 1.
 DR InterPro; IPR006650; A/AMP deam sub.
 DR InterPro; IPR001365; A/AMP deaminase.
 DR InterPro; IPR006330; A deaminase.
 DR Pfam; PF00962; A deaminase; 1.
 DR TIGRFAMs; TIGR01430; aden deam; 1.
 DR PROSITE; PS00485; A_DEAMINASE; FALSE NEG.
 KW Hydroxylase; Nucleotide metabolism; Complete proteome.
 FT ACT SITE 200 200 POTENTIAL.
 FT ACT SITE 248 248 POTENTIAL.
 FT ACT SITE 281 281 POTENTIAL.
 FT ACT SITE 282 282 POTENTIAL.
 SQ SEQUENCE 334 AA; 38023 MW; FC386C9C80DC789B CRC64;

Query Match 4.8%; Score 149; DB 1; Length 334;
 Best Local Similarity 27.3%; Pred. No. 0.0014;
 Matches 65; Conservative 39; Mismatches 106; Indels 28; Gaps 11;

QY 314 EYTDPSHPQLHVFLKQV--GLD-LVDDSKPERRPT-----KHMPTPEQWTVNFNP 364
 DB 100 EIRPAPPQHTQEDLNENDVVEAALQDGESKLGHNSNLILCSLRHDPV-ERSIDLVL 158

QY 365 AFSYAYIYCYANL-----FTLNKLRSGKM--TTIKFRPHAGEAGDVHDLAATFLCH 415
 DB 159 ANSYNEGVCAVDLAGNESDFPPELHKEAFDLAYONGIKITIHAGETGAENILKSIKLH 218

QY 416 --NTSHGNLRKS-PVLOYLYLQIGLAMSPLSN--NSLFLDYHRNPFPFFORGLNVS 470
 DB 219 ADRIHGIFAYKSEELQYVIE--NQVLEMCPSKSVNDKAVKNYKMHFKYFDLGVKVT 277

QY 471 LSTDPLQIHLTKRPLVEEYSIAASLWKLSCDCEIARNVYQSGFSHA 528
 DB 278 LNTDRT---VSNVSLVDEYLNLANIPDFGIEIKTVIRNGISASFATFEFKVLKK 332

RESULT 14
 ADD1_STRCO STANDARD; PRT; 387 AA.

AC Q867J7.
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable adenosine deaminase 1 (EC 3.5.4.4) (Adenosine aminohydrolase
 DE 1).
 GN ADD1 OR SC05662 OR SC6A9.05.
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953;
 RA Bentley S.D., Chater K.P., Cerdeno-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RA "Complete genome sequence of the model actinomycete Streptomyces
 RT coelicolor A3(2)."
 RL Nature 417:141-147(2002).
 CC -!- CATALYTIC ACTIVITY: Adenosine + H(2)O = inosine + NH(3).
 CC -!- SIMILARITY: Belongs to the adenosine and AMP deaminases family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; AL031035; CAA19890.1; -;
 DR PIR; T35436; T35436.

DR HAMAP; MF 00540; -; 1.
 DR InterPro; IPR006650; A/AMP deam sub.
 DR InterPro; IPR001385; A/AMP deaminase.
 DR InterPro; IPR006330; A deaminase.
 DR Pfam; PF00962; A deaminase; 1.

DR TIGRfams; TIGR01430; aden deam; 1.
 DR PROSITE; PS00485; A DEAMINASE; FALSE NEG.

KW Hydrolase; Nucleotide metabolism; Complete proteome.

FT ACT_SITE 244 244 POTENTIAL.
 FT ACT_SITE 292 292 POTENTIAL.
 FT ACT_SITE 325 325 POTENTIAL.
 FT ACT_SITE 326 326 POTENTIAL.

SQ SEQUENCE 387 AA; 42429 MW; 660F2621C115EA5A CRC64;

Query Match 4.5%; Score 138.5; DB 1; Length 387;
 Best Local Similarity 22.1%; Pred. No. 0.0099;
 Matches 95; Conservative 53; Mismatches 157; Indels 125; Gaps 19;

QY 108 EFLAQKTAP---HRDFYVNRKVDTHVHSACMNOKHL-----LRFIKSKLRKEPDEWVIFR 160

DB 37 EHLVDPDVDFDLHAFAGPKAELHVVHVSASPRIVSELAAKHADSKVPTDPEALV---- 93

QY 161 DGYMTLKEVFESLDLTGYDLNVDLDVHADKSTPHRFDKNLKYNPCQSQRLREIFLKQ 220

DB 94 --DYFTFTDFAHFDV--YLSVVDLIRTPED----- 120

QY 221 DNLQRFALUTKQVFSDLASKYQMAEYRISYGRKSEWDQLASWVNN-----EL 274

DB 121 -----VRLTYEYVARDMARQVRYAELTIPTFSSTRGIDEGAFMDAIEDAKAAEA 172

QY 275 HSGNVV-WLVQIPRLYVNYKEMGIVTSFQNLNDNIFVLFVETIDPASHPQLHVFLKQV 333

DB 173 EFTVURWCFDIP-----GEAGLESA-----ETARLAT----- 201

QY 334 GLDLVDESKPERRPTKHMPTPQWTVNFPNPAFSYYAYCYANLFTLNKLRKSGMTTIK 393

DB 202 ----DRLRPEGLVSFGLGPE--IGVARPQPKYFDRAIA-----AGLHSV- 242

QY 394 FRPHAGEAGDVHL--AATFLCHNISHGINLRKSPVLQVLYLGO--TGLAMSPLSNNS 449

DB 243 --PHAGETTPQPTVWEALIDLRAERIGHGTSSAQDPKL--LAHLAERRIPLEVCPPTSNTA 298

QY 450 L-----FLDYHRNPFPPTFFQGLNVLSTDDPLQIHLTKEPLVVEYSIAASLWKLSSCDLC 505

DB 299 TRAVRTLDEH--PIKEFVRAGVPTVINSDDP-----PMFGTDLNNEYAVARLLGLDERGLA 353

QY 506 ETARNVYQS 515

DB 354 DLAKNGVEAS 363

RESULT 15

ADD_ECOL6

ID ADD_ECOL6 STANDARD; PRT; 333 AA.

AC Q8FH99;

DT 15-SEP-2003 (Rel. 42, Created)

DT 15-SEP-2003 (Rel. 42, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Adenosine deaminase (EC 3.5.4.4) (Adenosine aminohydrolase).

GN ADD OR C2015.

OS Escherichia coli O6.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.

OX NCBI_TaxID=217992;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=O6:H1 / CFT073 / ATCC 700928;

RX MEDLINE=22388234; PubMed=12471157;

RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,

RA Raeko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,

RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,

RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;

RT "Extensive mosaic structure revealed by the complete genome sequence

of uropathogenic Escherichia coli.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).

CC -!- CATALYTIC ACTIVITY: Adenosine + H(2)O = inosine + NH(3).

CC -!- SIMILARITY: Belongs to the adenosine and AMP deaminases family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AE016761; AAN80475.1; -;

DR HAMAP; MF 00540; -; 1.

DR Pfam; PF00962; A deaminase; 1.

DR TIGRfams; TIGR01430; aden deam; 1.

DR PROSITE; PS00485; A DEAMINASE; 1.

KW Hydrolase; Nucleotide metabolism; Complete proteome.

FT ACT_SITE 197 197 POTENTIAL.

FT ACT_SITE 245 245 POTENTIAL.

FT ACT_SITE 278 278 POTENTIAL.

FT ACT_SITE 279 279 POTENTIAL.

SQ SEQUENCE 333 AA; 36360 MW; A89FCBC00F5D3F47 CRC64;

Query Match 3.9%; Score 121.5; DB 1; Length 333;

Best Local Similarity 24.7%; Pred. No. 0.13; Mismatches 41; Gaps 7;

Matches 47; Conservative 27; Indels 75; Indels 41; Gaps 7;

QY 318 DPASHPQLHVFL---KQVGLDLDVDESKPERRPTKHMPTPEQWTVNFPNPAFSYYAYCY 374

DB 146 EAACQOELEAFLAHRDQITALDLAGDE-----LGFPGLSLSHF-- 184

QY 375 ANLFTLNKLRKSGMTTIKFRPHAGEAGDVHL--AATFLCHNISHGINLRKSPVLQYL 432

DB 185 -----NRARDAGWHITV---HAGEAAGPESIQWAIRELGAERIGHGVKATIEDRALMDF 234

QY 433 YVLGIGLAMSPLSN--NSLFLDYHRNPFPFFQGLNVLSTDDPLQIHLTKEPLVVEY 490

DB 235 LAEQQIGIESCLTSNIQTSTVADLAHPLKLTLEHGIRASINTDDP----GVQGVDIHEY 291

QY 491 SIAASLWKLKLS 500

DB 292 TVAAPAAGLS 301

Search completed: December 15, 2003, 13:27:14

Job time : 20 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 13:24:28 ; Search time 40 Seconds
(without alignments)
3728.860 Million cell updates/sec

Title: US-10-019-633-2

Perfect score: 3100

Sequence: 1 PRRVAPWEKEVINDPCTPK.....NEMKLVYSDNEILLIPDEL 578

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL.23.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2644.5	85.3	839	10 O80452	O80452 arabidopsis
2	1798	58.0	743	5 Q9NGX0	Q9NGX0 dictyosteli
3	1687.5	54.4	798	11 Q9DBT5	Q9DBT5 mus musculus
4	1680.5	54.2	798	4 Q961A1	Q961A1 homo sapien
5	1641	52.9	1008	3 Q9P3N4	Q9P3N4 neurospora
6	1638	52.8	652	5 Q9VY76	Q9VY76 drosophila
7	1638	52.8	707	5 Q961O7	Q961O7 drosophila
8	1569	50.6	865	5 Q18463	Q18463 caenorhabdi
9	1569	50.6	889	5 Q95ZX5	Q95ZX5 caenorhabdi
10	1537	49.6	766	11 Q8CFR4	Q8CFR4 mus musculus
11	1424.5	46.0	481	11 Q922J7	Q922J7 mus musculus
12	1325	42.7	672	5 Q8IE53	Q8IE53 plasmodium
13	1060.5	34.2	1655	5 Q9GRM4	Q9GRM4 leishmania
14	954.5	30.8	2093	5 Q9XZY8	Q9XZY8 leishmania
15	901.5	29.1	286	11 Q91Y12	Q91Y12 mus musculus
16	581	18.7	288	5 Q9GRR7	Q9GRR7 leishmania

17	166	5.4	553	5	Q26642	Q26642 sarcophaga
18	150	4.8	331	16	Q8EBD4	Q8EBD4 shewanella
19	147.5	4.8	543	5	Q8SP8	Q8SP8 dictyosteli
20	141	4.5	364	3	Q8X1T6	Q8X1T6 emericeila
21	138	4.5	339	3	Q9P6J8	Q9P6J8 schizosacch
22	136	4.4	488	5	Q813V7	Q813V7 plasmodium
23	134	4.3	553	5	Q9VVK5	Q9VVK5 drosophila
24	132.5	4.3	282	2	Q8LOX7	Q8LOX7 lactobacill
25	132.5	4.3	502	5	Q95WT8	Q95WT8 culex quing
26	132	4.3	591	3	Q8N1Z8	Q8N1Z8 neurospora
27	130	4.2	501	5	Q9VFS0	Q9VFS0 drosophila
28	130	4.2	506	5	Q8WS97	Q8WS97 drosophila
29	121.5	3.9	332	12	Q91GY8	Q91GY8 orf virus.
30	121.5	3.9	333	16	Q8FH99	Q8FH99 escherichia
31	121.5	3.9	334	12	Q9YRR3	Q9YRR3 orf virus
32	121.5	3.9	334	12	Q91GY9	Q91GY9 orf virus.
33	120	3.9	4717	3	Q94248	Q94248 schizosacch
34	120	3.9	5890	5	Q81K84	Q81K84 plasmodium
35	119.5	3.9	2543	5	Q81BR8	Q81BR8 plasmodium
36	118.5	3.8	340	12	Q91OT9	Q91OT9 orf virus.
37	118.5	3.8	340	12	Q91GZ0	Q91GZ0 orf virus.
38	118	3.8	506	5	Q9U7C5	Q9U7C5 glossina mo
39	117	3.8	39	6	Q9TUK2	Q9TUK2 sus scrofa
40	117	3.8	442	16	Q8EZR9	Q8EZR9 leptospira
41	116	3.7	1320	5	Q81LJ3	Q81LJ3 plasmodium
42	116	3.7	3347	5	Q81EAL	Q81EAL plasmodium
43	116	3.7	4638	5	Q81K96	Q81K96 plasmodium
44	115.5	3.7	530	5	Q8T9T6	Q8T9T6 aedes aegypt
45	114.5	3.7	435	16	Q98Q33	Q98Q33 mycoplasma

ALIGNMENTS

RESULT 1

O80452 PRELIMINARY; PRT; 839 AA.
 ID O80452
 AC O80452; Q93ZR9;
 DC 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Putative AMP deaminase.
 GN ATG38280 OR F16M14.21/ATG38280.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_taxid=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Rounsley S.D., Kaul S., Lin X., Ketchum K.A., Crosby M.L.,
 RA Brandon R.C., Sykes S.M., Mason T.M., Kerlavage A.R., Adams M.D.,
 RA Somerville C.R., Venter J.C.;
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Lin X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Town C.D., Kaul S.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Banh J., Banno F., Dale J.M., Goldsmith A.D., Lee J.M.,
 RA Onodera C.S., Quach H.L., Tang C., Toriumi M., Yamamura Y., Yu G.,
 RA Yu S., Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,
 RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C.,
 RA Koesema E., Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M.,
 RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,


```

DT 01-JUN-2001 (T-EMBLrel. 17, Created)
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE 1200014F01RIK
GN 1200014F01RIK
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuoka H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Cariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaeerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Suzuki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
DR EMBL; AK004759; BAB23540.1; -
DR MGD; MGI:1921388; 1200014F01RIK.
DR InterPro; IPR001365; A/AMP deaminase.
DR InterPro; IPR006650; A/AMP_deam_sub.
DR InterPro; IPR006329; AMP_deaminase.
DR Pfam; PF00962; A_deaminase; 1.
DR TIGRFAMs; TIGR01429; AMP deaminase; 1.
DR PROSITE; PS00485; A_DEAMINASE; 1.
DR PROSITE; PS00485; A_DEAMINASE; 1.
SQ SEQUENCE 798 AA; 92070 MW; 2BC4F37E4006C7D5 CRC64;

Query Match 54.4%; Score 1687.5; DB 11; Length 798;
Best Local Similarity 54.2%; Pred. No. 1.2e-127;
Matches 318; Conservative 84; Mismatches 162; Indels 23; Gaps 6;

QY 1 PRVRVAPWEKEVINDPCTPKPNPNFTYVPEPKSEHVFTQVGVHVA----DKDCTES 56
DB 220 PALEQHPYEH-----CEPSAMPGDLGL-----GLRMVGVVHVYTRRDPDHCPEV 265
QY 57 IYPVADATTFDTLHYLRVTAAGNTRTVCHNRLNLEHKFKFHLMLNADREFLAQKTAP 116
DB 266 ELFPYDLQEFVADVNVLMALIINGIKSCFYRLQVLSKSFQVHLLNEMKELAAQKVP 325
QY 117 HRDYNVRKVDTHVHSAACMNQKHLRFIKSKLRKEPEDEWVIFRDGYMTLKEVFESLDL 176
DB 326 HRDYNVRKVDTHVHSAACMNQKHLRFIKSKLRKEPEDEWVIFRDGYMTLKEVFESLDL 176
QY 177 TGYDLNVLLDVAHDKSTFHRDKNLKNPCQSRLEIFLKQDNLIQGRFLAELTKQV 236
DB 386 TAYDLSVDTLDVHADNTRFHRDKNLKNPCQSRLEIFLKQDNLIQGRFLAELTKQV 236
QY 237 FSDLASAKYQMAEYRISYGRKQSEWDLASWIVNNELHSGNVVWLQVIRLYNYVYKEMG 296
DB 446 MADLESKYQNAELRLSYIGRDEWDLARWAVNKHSPNVRVLVQVPRFLDVRTKG 505
QY 297 IVTSFQNLNDFVLPFEVITDPASHPOLHVLKQVGLDLDVDDSKPERRP-TKHMTPT 355
DB 506 QLAFQEMLENIFLPLFEATVHPASHPELHLFLEHVDGDFSDVDDSKPENHVNLESPLP 565
QY 356 EQWTVNFPAPSYAYCYANLFTLNKLRSGKMTTIKFRPHAGAGDVHUAATFLFLCH 415
DB 565 EQWTVNFPAPSYAYCYANLFTLNKLRSGKMTTIKFRPHAGAGDVHUAATFLFLCH 415

Db 566 EAWTEEDNPPYAYLYYTFANMAMNLHRRQGRFHTFVLRPHCGEAGDIHHLVSAFMLAE 625
QY 416 NISHGINLRKSPVLQYLYLQIGLAMSPLSNNSLFLDYHRNPFTFFQFQGLNLSLSTDD 475
DB 626 NISHGLLARKAPVLQYLYLAQIGLAMSPLSNNSLFLSYHRNPPLPEYLSRGLWSLSTDD 685
QY 476 PLQHLTKPELVEYSIAASLWKLSSCDLCEIARNSVYQSGFSHALKAHWIKGNVFKRGP 535
DB 686 PLQHFHTKEPLMEYSIATQVWKLSSCDLCEIARNSVYQSGFSHALKAHWIKGNVFKRGP 745
QY 536 AGNDIHRNTPVHIRVQPREMIWRNEMKLVYS--DNEIL--IPDELDEL 578
DB 746 EGNDIRTNVEDIRVGRYETLCOELALITQAVOSEMLETIPEEVGI 792

RESULT 4
Q96IAL PRELIMINARY; PRT; 798 AA.
AC Q96IAL;
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Adenosine monophosphate deaminase 2 (Isoform L).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC007711; AAH07711.1; -
DR InterPro; IPR001365; A/AMP deaminase.
DR InterPro; IPR006650; A/AMP_deam_sub.
DR InterPro; IPR006329; AMP_deaminase.
DR Pfam; PF00962; A_deaminase; 1.
DR TIGRFAMs; TIGR01429; AMP deaminase; 1.
DR PROSITE; PS00485; A_DEAMINASE; 1.
DR PROSITE; PS00485; A_DEAMINASE; 1.
SQ SEQUENCE 798 AA; 92070 MW; A31876E9F8C93EB CRC64;

Query Match 54.2%; Score 1680.5; DB 4; Length 798;
Best Local Similarity 54.3%; Pred. No. 4.3e-127;
Matches 317; Conservative 82; Mismatches 162; Indels 23; Gaps 6;

QY 1 PRVRVAPWEKEVINDPCTPKPNPNFTYVPEPKSEHVFTQVGVHVA----DKDCTES 56
DB 219 PALEQHPYEH-----CEPSTMPGDLGL-----GLRMVGVVHVYTRRDPDHCSEV 264
QY 57 IYPVADATTFDTLHYLRVTAAGNTRTVCHNRLNLEHKFKFHLMLNADREFLAQKTAP 116
DB 265 ELFPYDLQEFVADVNVLMALIINGIKSCFYRRLQVLSKSFQVHLLNEMKELAAQKVP 324
QY 117 HRDYNVRKVDTHVHSAACMNQKHLRFIKSKLRKEPEDEWVIFRDGYMTLKEVFESLDL 176
DB 325 HRDYNVRKVDTHVHSAACMNQKHLRFIKSKLRKEPEDEWVIFRDGYMTLKEVFESLDL 176
QY 177 TGYDLNVLLDVAHDKSTFHRDKNLKNPCQSRLEIFLKQDNLIQGRFLAELTKQV 236
DB 385 TAYDLSVDTLDVHADNTRFHRDKNLKNPCQSRLEIFLKQDNLIQGRFLAELTKQV 444
QY 237 FSDLASAKYQMAEYRISYGRKQSEWDLASWIVNNELHSGNVVWLQVIRLYNYVYKEMG 296
DB 445 MSDLSESKYQNAELRLSYIGRDEWDLARWAVNKHSPNVRVLVQVPRFLDVRTKG 504
QY 297 IVTSFQNLNDFVLPFEVITDPASHPOLHVLKQVGLDLDVDDSKPERRP-TKHMTPT 355
DB 505 QLAFQEMLENIFLPLFEATVHPASHPELHLFLEHVDGDFSDVDDSKPENHVNLESPLP 564
QY 356 EQWTVNFPAPSYAYCYANLFTLNKLRSGKMTTIKFRPHAGAGDVHUAATFLFLCH 415
DB 565 EAWTEEDNPPYAYLYYTFANMAMNLHRRQGRFHTFVLRPHCGEAGDIHHLVSAFMLAE 624

```


SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champagne M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dreanek D., Farfan D.,
RA Ferreria S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirkas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RA "Sequencing of Drosophila melanogaster genome."
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RA "Annotation of Drosophila melanogaster genome."
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[4]
SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[5]
SEQUENCE FROM N.A.
RA FlyBase;
RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AS003493; AAF49329.2; -
DR FlyBase; Fgn0052626; CG32626.
DR InterPro; IPR001365; A/AMP_deaminase.
DR Pfam; PF00962; A_deaminase; 1.
DR PROSITE; PS00485; A_DEAMINASE; 1.
SQ SEQUENCE 652 AA; 76054 MW; 846D8012F2E277E2 CRC64;
Query Match 52.8%; Score 1638; DB 5; Length 652;
Best Local Similarity 53.9%; Pred. No. 8.e-124;
Matches 305; Conservative 87; Mismatches 152; Indels 22; Gaps 5;
7 PNEKEVINDPCTKPNPNFTVVPKSEHVQTVGVHLYVA-DKDCETESYPVADATT 65
96 PNWVEP-----PNDEDFKTKP-----LNGVFIHYENDESSEIKYEYPMDSQ 137
66 FFTDLHYILRVTAAGNTRVCHNRLNLLHKKFHLMLNADREFLAQKTAHPDRFVNVK 125
138 FVNDQVQVMCMNLTADGLPKSCFYRRLCYLSSKQVMHVLNELLRLAQAQVHPDRPYNTRK 197
126 VTHVHSAQMKQKLLRFKIKLKEPDEVVIFRDGYTYMLKVFESLDLTGYLDNVDL 185
198 VDTTHAASQMKQKLLRFKIKLKEPDEVVIFRDGYTYMLKVFESLDLTGYLDNVDL 257
186 LDVHADKSTFHFEDKENLYNFCGOSRLREIFLKQDNLTQGFLELTQVPSDLSASKY 245
258 LDVHADRNTHFRDFKFNKYNPIGESRLREVEFLKTDNYLNGKYFAQIIKEVAFDEESKY 317
246 QMAEYRISYIGKQSEWDQLASVNVNLSHNVVWLVOIPLRYNXYKEMGVITSPONLL 305
318 QNAELRLSYIGKSPDEWYKLAWADNDVSYNIRLWLIQPLRFDIFKSNQMKQFQELL 377
306 DNIPLVPLEFTVITDPASHPOLHFLKQVGLDLDVDDSKPERP--TKHMTPEQWNTVFN 363
378 NNIFLPLFEATAPSKHPHFLQVIGVIGFSDVDDSKPE-NPLFDNDVPRPEWNTYEN 436
364 PAFSYAYCYANLFTNLKLRSGKMTTKIFRPHAGEAGDVHDLAATFLCLNISHGNL 423

Db 437 PPAYIYYIYVYANMTVLNKFQSRNMNTFVLRPHCGEAGPVQHLVCGFLMAENISHGLLL 496
Qy 424 RKSPLVQLYLYIGQIGLAMSPLSNNSPLDHYRNPFTFFQFGLNVLSLSTDDPLQIHLTK 483
Db 497 RKPVLQVLYLYITQIGIAMSPLSNNSPLFNHRNPPLPEYLRGLIISLSTDDPLQFHFTK 556
Qy 484 EPLVEYSIAASLWKLSSCDLCEIARNVSYGSGSHALKAHWIKVKNYKFGPAGNDIIRT 543
Db 557 EPLMEYSIAAQVWKLSSCDLCEIARNVSMGSPHAIKQOQMLGPIYYEDGIMGNDIIRT 616
Qy 544 NVPHIRVQFREMWRNEMKLVYSNE 569
Db 617 NVPEIRVAYRYETLDELSENFVKVQ 642
RESULT 7
Q96107 PRELIMINARY; PRT; 707 AA.
ID Q961Q7;
AC Q961Q7;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE GH10492p (CG32626-PC)
GN CG32626 OR CG11058 OR CG11065 OR CG15762.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RA Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
EX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abail J.F., Agbayani A., An H.J., Andrews-Frankoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.I., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merklov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wasserman D.A., Weinstock G.M., Weissbach J.,

RA Williams S.M., WoodageT, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.:
 RT "The genome sequence of *Drosophila melanogaster*;"
 RL Science 287:2185-2195(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Celnik S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Anantides P.G., Brandon R.C., Rogers Y.,
 RA Ebanon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
 RA Ferrera S., Frise E., Galle R.F., Gaig N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
 RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
 RT "Sequencing of *Drosophila melanogaster* genome;"
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celnik S.E.,
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
 RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,
 RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
 RT "Annotation of *Drosophila melanogaster* genome;"
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Adams M.D., Celnik S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
 RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RA FlyBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY051429; AAK92853.1; -;
 DR EMBL; AE003493; AAN09337.1; -;
 DR FlyBase; FBgn0052626; CG32626.
 DR InterPro; IPR001385; A/AMP_deaminase.
 DR InterPro; IPR006650; A/AMP_deam sub.
 DR InterPro; IPR006329; AMP_deaminase.
 DR Pfam; PF00962; A_deaminase; 1.
 DR TIGRFAMs; TIGR01429; AMP_deaminase; 1.
 DR PROSITE; PS00485; A_DEAMINASE; 1.
 SQ SEQUENCE 707 AA; 82035 MW; 0FE03C647022051A CRC64;

Query Match 52.8%; Score 1638; DB 5; Length 707;
 Best Local Similarity 53.9%; Pred. No. 9.9e-124;
 Matches 305; Conservative 87; Mismatches 152; Indels 22; Gaps 5;

QY 7 PWEKEVINDCTPKPNPNPTTYPEPKSEHVFTVDGVHIVA-DKDCIESIPVADATT 65
 151 PNWVEP-----PNDEDPKFKP-----LNGVPHIYENDDESSEIKYEYDMSQ 192
 66 FFTDLHVLRTVTAAGNTRTCHNRLNLEHKKFHLMLNADREFLAQKTAAPHDRFYNRK 125
 193 FVNDQMVMCMNIADGPKSFYRRLCYLVSSKYQWVLLNELRLAQAQVPHDRDFNTRK 252
 126 VDTVHHSACMOKHLLRFIKSKLRKEPDEWIFRDGTVMTLKEVFSLDLTGYDLNVDL 185
 253 VDTTHAASCMWQKHLRFIKTKLKNANEVVTVTNGQQMTLAQVFQSMNLTYYDLTVD 312
 186 LDVHADKSTHRTFDKFNKLNKPNCGSRLRIFLQKODNLIOGRFLAELTQVDFDLSASKY 245
 313 LDVHADRNTHRTFDKFNKSNYPNGESRLREVFLKTDNYLNGKYFAQIIEKVAFDLESKY 372

QY 246 QMAEYRISYGRKQSEWDLASWIVNNELHSGNVVWLVOIPRLNXYKEMGIVTSFQNL 305
 DB 373 QNAELRLSIYCKSPDEWYKLAKWAINDVSYSNIRLWLIQIPRLDFIKSNKMKSFQIL 432
 QY 306 DNIPLPEEVTIDPASHPOLHVLKQVVGDLVDDESKPERRP--TKHMPPTQOWTNVFN 363
 DB 433 NNIPLPEATARPSKGPHELRFLOYVIGFVDSVDDESKPE-NPLFDNDVPRPEEWTYEN 491
 QY 364 PAFSYAYCYANLFTLNKLRKESGWTIKFRPHAGEAGVDVHLAATAFLCHNISHGINL 423
 DB 492 PPAYIYYWYANVTNLNFKQSRNMNTFVLRPHCGEAGPVQHLVCGFLMAENISHGLLL 551
 QY 424 RKSPLQYLYLQIGQIGLAMSPLSNNSFLDYHRNPFFTFQFQGLNVLSLSTDDPLQIH 483
 DB 552 RKVPVLQYLYLTQIGIAMSPLSNNSFLNHRNPFLPEYLARGLIISLSTDDPLQHFHTK 611
 QY 484 EPLVEYSIAASLWKLSSCDLCETARNVYOSGSHALKAHWIGKNYKFRGPAGNDIHT 543
 DB 612 EPLMEYSIAAQVWKLSSCDMCELARNVSMGSPFHAIKQWMLGPIYYEDGIMGNDITRT 671
 QY 544 NVPHIRVOFREMIMWRNEMKLVYSNNE 569
 DB 672 NVPEIRVAYRYETLLDELSNIFKNQ 697

RESULT 8

Q18463
 ID Q18463 PRELIMINARY; PRT; 865 AA.
 AC Q18463;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical 99.5 kDa protein.
 GN C34F11.3
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode *C. elegans*: a platform for
 RT investigating biology. The *C. elegans* Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Bentley D.;
 RT "The sequence of *C. elegans* cosmid C34F11.";
 RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RT "Direct Submission.";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U46753; AAA85762.1; -;
 DR WormPep; C34F11.3a; CE04164.
 DR InterPro; IPR001365; A/AMP_deaminase.
 DR InterPro; IPR006650; A/AMP_deam sub.
 DR InterPro; IPR006329; AMP_deaminase.
 DR Pfam; PF00962; A_deaminase; 1.
 DR TIGRFAMs; TIGR01429; AMP_deaminase; 1.
 DR PROSITE; PS00485; A_DEAMINASE; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 865 AA; 99499 MW; EA22B7D62D12EDFD CRC64;

Query Match 50.6%; Score 1569; DB 5; Length 865;
 Best Local Similarity 53.5%; Pred. No. 5e-118;
 Matches 303; Conservative 88; Mismatches 155; Indels 20; Gaps 7;

DR EMBL; BC040366; AAH40366.1; --
SQ SEQUENCE 766 AA; 88652 MW; 94991133B18B8AA5 CRC64;

Query Match 49.8%; Score 1537; DB 11; Length 766;
Best Local Similarity 51.8%; Pred. No. 1.6e-115;
Matches 288; Conservative 91; Mismatches 163; Indels 14; Gaps 4;

QY 19 PKENPNPFYVPEPKS-EHVFQTVGVGVIHVYADKQCTE---SIVPVADATFTFFDLHYI 73
DB 201 PUPQEDPYCLDDAPPNIGLVLRMHGGVLFYDQNTMLERQEPHSLPYDPLETYIYVDMSHI 260

QY 74 LRVTAAGNTRTCHNRLNLEHKEFHLMLNADREFLAQTAHPRDFYNNRVKVDTHVHRS 133
DB 261 LALITDGPTKYCHRELNFLESKFSLEHLMENSEFELKSNPRDFYNNRVKVDTHHAA 320

QY 134 ACNQKHLLRFKSKURKSPDEVVIIPRDGYMTLKEVPESLDLTGVDLNVLDLVDHAKS 193
DB 321 ACNQKHLLRFKIKHTYTQTPDRTVAEKLGRKITLRQVDFSLHMDPYDLTVDSLDVHAGRQ 380

QY 194 THERFDKFNKYNPCQSQSLREIFLKQDNLIOGRFLAELTKQVFSLSASKYQMAEYRIS 253
DB 381 THERFDKFNKYNPVGASELRDLDYLTENYLGGEYFARMVKEVARELEDSKYQSEPLUS 440

QY 254 IYGRKQSEVDQLASWIVNNELHSGNVVWLQVLPRLYNNYKEMGIVTSFQNLLDNIFVPLF 313
DB 441 IYGRSPKEWSSLARWFIQHKVYSPNRRWIIQVPIYDIFRSKLLPNFGKMLENIFLPLF 500

QY 314 EVTIDPASHPOLHVLKQVVGDLVDDESKPERPPTKHM-----PTPEQWTVNVPAPSY 368
DB 501 KATINPDQHRELHLFLKYVTGFSVDDESK---HSDHMFSDKSPSPDLMTSEQNPPSY 556

QY 369 YAYCYANLFTLNKRESKGMTIKERPHAGEAGDVHLLAATFLCHNLSHGINKRSPV 428
DB 557 YLYTYANIMVNLNRRERGLSTFLPRPHCGEAGSITHLVSAFLTADNISHGLLLKSPV 616

QY 429 LOVLYLGOIGLAMSPLSNLSFLDYHRNPFPTFFQRLGNVSLSTDDPLQIHLTKPELVE 488
DB 617 LOVLYLAQIPIAMSPLSNLSFLEYSKNPLREFLHKLHLSLSTDDPMQPHYTKALME 676

QY 489 EYSIASLWKLSCDICEIARNVYQSGFHALKAHWIGKNYKPKRPGAGNDIHRNTVPHI 548
DB 677 EYVIAAQVWKLSTCDLCEIARNVSLQSLSHQKQKFLGQNYKKEGPEGNDIRKTNVAQI 736

QY 549 RVQFREMIRWNEKLV 564
DB 737 RMAFRYETLCNELSFL 752

RESULT 11
Q922J7 PRELIMINARY; PRT; 481 AA.

AC Q922J7;
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Unknown (Protein for IMAGE:3591312) (Fragment).
GN AMPD3.

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL EMBL; BC007183; AAH07183.1; --
DR MGD; MGI:1096344; Ampd3.
DR InterPro; IPR001365; A/AMP deaminase.
DR InterPro; IPR006650; A/AMP deam_sub.
DR InterPro; IPR006329; AMP deaminase.
DR Pfam; PF00962; A_deaminase; 1.
DR TIGRFAMs; TIGR01429; AMP deaminase; 1.
DR PROSITE; PS00485; A_DEAMINASE; 1.

FT NON TER 1
SQ SEQUENCE 481 AA; 56376 MW; B6DF4F177F105A33 CRC64;

Query Match 46.0%; Score 1424.5; DB 11; Length 481;
Best Local Similarity 55.5%; Pred. No. 1.1e-106;
Matches 261; Conservative 77; Mismatches 123; Indels 9; Gaps 2;

QY 100 HMLNADREFLAQTAHPRDFYNNRVKVDTHVHHSACNMQKHLLRFKSKURKPEDEVVIF 159
DB 2 HMLNEMSEFELKSNPRDFYNNRVKVDTHHAAACNMQKHLLRFKIKHTYTQTPDRTVAE 61

QY 160 RDGYMTLKEVPESLDLTGVDLNVLDLVDHAKDSTFHRFDKFNKYNPCQSQSLREIFLK 219
DB 62 KLGRKITLRQVDFSLHMDPYDLTVDSLDVHAGRQTFHRFDKFNKYNPVGASELRDLYLK 121

QY 220 QDNLIQGRFLAELTKQVFSLSASKYQMAEYRISYGRKQSEVDQLASWIVNNELHSGNV 279
DB 122 TENYLGGEYFARMVKEVARELEDSKYQSEPLUSIYGRSPKEWSSLARWFIQHKVYSPN 181

QY 280 VMLQVLPRLYNNYKEMGIVTSFQNLLDNIFVPLPEVTIDPASHPOLHVLKQVVGDLVD 339
DB 182 RWIIQVPIYDIFRSKLLPNFGKMLENIFLPLFKATINPDQHRELHLFLKYVTGFSVD 241

QY 340 DESKPERPPTKHM-----PTPEQWTVNVPAPSYAYCYANLFTLNKRESKGMTIKF 394
DB 242 DESK---HSDHMFSDKSPSPDLMTSEQNPPSYLYTYANIMVNLNRRERGLSTFL 297

QY 395 RPHAGEAGDVHLLAATFLCHNLSHGINKRSPVLYLYLGOIGLAMSPLSNLSFLDY 454
DB 298 RPHCGEAGSITHLVSAFLTADNISHGLLLKSPVLYLYLAQIPIAMSPLSNLSFLEY 357

QY 455 HRNPPTFFQRLGNVSLSTDDPLQIHLTKPELVEYSIAASLWKLSCDICEIARNVYQ 514
DB 358 SKNPUREFLHKLHLSLSTDDPMQPHYTKALMEYVIAAQVWKLSTCDLCEIARNV 417

QY 515 SGFSHALKAHWIGKNYKPKRPGAGNDIHRNTVPHIRVQFREMIRWNEKLV 564
DB 418 SGLSHQKQKFLGQNYKKEGPEGNDIRKTNVAQIRMAFRYETLCNELSFL 467

RESULT 12
Q9IE53 PRELIMINARY; PRT; 672 AA.

AC Q9IE53;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE AMP deaminase, putative (EC 3.5.4.6).
GN MAL13P1.146.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OC NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RA Harris B., Lennard N., Clark L., Line A., Barron A., Corton C.,
RA Berriman M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J.,
RA Ormond D., Sanders M., Hayes B., Hall S., Quail M., Barrell B.,
RL Submitted (SEP-2002) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AL844509; CAD52413.1; --
KW Hydrolase.

SQ SEQUENCE 672 AA; 79308 MW; F50A6CF4CEA62CCB CRC64;

Query Match 42.7%; Score 1325; DB 5; Length 672;
Best Local Similarity 47.3%; Pred. No. 1.9e-98;
Matches 268; Conservative 89; Mismatches 186; Indels 24; Gaps 8;

QY 8 WEKEVINDPCTPKPNPNPTFYVPEPKSEHVFQTVGVHIVYAD-----KD-CTES- 56
DB 109 YKSNIDDYSESSEPIYNPNVKILKNCNAFINFVDGIFVHWDPTDEGSSRDMCVESN 168

QY 57 -----IYPVADATFTFDLHYILRVTAAGNTRVCHNRLNLEHKEFHLMLNADREFLAQ 112
DB 169 KLANHRNTKASDYLSIQEIMNVVQDPACKSFCFQRLKYLEKPFDFHMFNGPLELSET 228

Best Local Similarity 37.5%; Pred. No. 8.3e-68;		Matches 169; Conservative 34; Mismatches 72; Indels 5; Gaps 3;	
Matches 206; Conservative 100; Mismatches 219; Indels 25; Gaps 8;			
QY 27 TVPEPKSHVFTVDGVI---HVVADKDCSTSIYPVADATFTFDLHYI---LRVTAAG 80			
Db 1249 SYLQFDEDEWKYKTEGIIIVPHEVHQIPRLQPDQMYHYTEFCNHVEEIRCLDNIIRV--- 1304			
QY 81 NRTVTCHNLNLEHKEFKFHLMLNADREF--LAQKTAPHRDFYNNVKVDVTHVHHSACNMQ 138			
Db 1305 --RDFALRQLLLEHFRKLUHAAVNHSHRELGSSTAARASHNRDFTQSTKVDNNIRMETGMTA 1362			
QY 139 KHLIRFIKSLRKEPDEVVIFRDGTY-MTLKEYVESLDLTGYDLNVLDDLVHADSKSTFHR 197			
Db 1363 RQLLAIFIVDKATHNGDDIVSHPKGEPQTLRQLLADLHITADSLTVDDLNVQAGATS--- 1419			
QY 198 FKFENLKNPCG-----QSLREIFLQDNLTQGRFLABLTQVFSDDLASKYQVAEYRI 252			
Db 1420 -SNGGAPQNPFASEGQCOQDELLTLTKTDNQMGRYFABLTKRTFEELSRDQHTFSESRL 1478			
QY 253 SIYGRKQSEMDOLASWIVNNELHSGNVWLVOIPRLYNNVYKEMGIVTSFQNLDDNIFVPL 312			
Db 1479 SVYGASAEWALLSHWFDTHGSSSHNQWVQPRIYSIRKAGRVASFAEYLEHVFEP 1538			
QY 313 FEVTIDPASHPOLHVELKQVGVGLDLDVDESKPERRPTKHMPTPEQWNTVNFPAFSYAYY 372			
Db 1539 WRISLHPNSDPRLFHEINHIAFDCEVEDERRPDVPLHLMRSPHEWTTEDPPYNYLYH 1598			
QY 373 CVANLFTLNKLRSGKMTTIKFRPHAGEAGDVHDLAATFLCHNISHGINLRKSPVLQYL 432			
Db 1599 LYANLRLSLNCFQRRRFSFVSFRPSCGEAGGVHLLIGGLAQSVNYGVRLADSAFLOYL 1658			
QY 433 YLGOGLGLAMSPLSNNS-LFLDYHRNPPTFFFORGLNVSLSTDDPQLOIHLTKPELVEEYS 491			
Db 1659 FYLAQIGVTLSPUSNNKQLNLYHNPPFOFFRGRURVSLGTDSPLLYHHTQEPLEEYS 1718			
QY 492 IAASLWKLSSCDLCEIARNVYQSGFSHALKAHWIGKNYFKRGPAGNDIHRNTVPHIRVQ 551			
Db 1719 IASKIWKLSPNDLSEVARNVLLSNFSLRKEEKLGMHFLSSSAGNDVAKTHLSDRVYA 1778			
QY 552 FREMIWRNEM 561			
Db 1779 YRFEAYHTEV 1788			
RESULT 15			
Q91Y12	PRELIMINARY;	PRT;	286 AA.
AC Q91Y12;			
DT 01-DEC-2001 (T-EMBLrel. 19, Created)			
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)			
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)			
DE Hypothetical protein.			
GN 1200014F01RIK.			
OS Mus musculus (Mouse).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX NCBI_TaxID=10090;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC STRAIN=FVB/N;			
RA Strausberg R.;			
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.			
DR EMBL; BC016662; AA016662.2; -			
DR MGD; MGI:1921389; 1200014F01RIK.			
DR InterPro; IPR001365; A/AMP deaminase.			
DR InterPro; IPR006650; A/AMP deam_sub.			
DR Pfam; PF00962; A deaminase; 1.			
DR PROSITE; PS00485; A DEAMINASE; 1.			
KW Hypothetical protein.			
SQ SEQUENCE 286 AA; 32677 MW; 43106743E178C39D CRC64;			
Query Match 29.1%; Score 901.5; DB 11; Length 286;			
Best Local Similarity 60.4%; Pred. No. 1e-64;			

Search completed: December 15, 2003, 13:28:07
Job time : 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 13:25:03 ; Search time 20 Seconds
(without alignments)
2779.275 Million cell updates/sec

Title: US-10-019-633-2
Perfect score: 3100
Sequence: 1 PRVRVAPWEKEVINDPCTPK.....NEWKLVSNDNEILIPDEL 578

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76: *
1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2624.5	84.7	600	T01259	AMP deaminase homo
2	1680.5	54.2	760	A44313	AMP deaminase (EC
3	1647	53.1	810	S49744	AMP deaminase (EC
4	1641	52.9	924	T50996	probable AMP deami
5	1617	52.2	846	T3261	amp deaminase - fi
6	1569	50.6	865	T15771	hypothetical prote
7	1549	50.0	774	S68147	AMP deaminase (EC
8	1549	50.0	776	S68146	AMP deaminase (EC
9	1505	48.5	747	A27366	AMP deaminase (EC
10	1497	48.3	747	I39444	AMP deaminase (EC
11	1186.5	38.3	605	S59996	AMP deaminase (EC
12	627	20.2	888	S50801	AMP deaminase homo
13	625	20.2	797	S44546	probable membrane
14	350	11.3	88	A37056	AMP deaminase (EC
15	149	4.8	334	G97269	adenosine deaminas
16	138.5	4.5	387	T35436	probable adenosine
17	121.5	3.9	339	T11785	adenosine deaminas
18	120.5	3.9	352	G86660	adenosine deaminas
19	120	3.9	471	T41581	hypothetical coile
20	117.5	3.8	333	A64919	adenosine deaminas
21	117.5	3.8	333	C90920	adenosine deaminas
22	117.5	3.8	333	H85768	adenosine deaminas
23	115.5	3.7	3071	T50345	vacuolar protein s
24	114.5	3.7	435	H90578	conserved hypothet
25	112.5	3.6	868	G71691	hypothetical prote
26	112.5	3.6	873	B86471	hypothetical prote
27	112	3.6	888	E82885	hypothetical prote
28	112	3.6	1384	A44396	P-type cation tran
29	111.5	3.6	316	H83625	probable adenosine

RESULT 1

T01259
AMP deaminase homolog F16M14.21 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 31-Mar-2003
C:Accession: T01259; A84803
R:Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L. euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 763-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: A84803
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-600 <ROU>
A:Cross-references: ENBL:AC003028; NID:g3335356; PID:g3335375
A:Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L. euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 763-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: A84803
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-600 <STO>
A:Cross-references: GB:AB002093; NID:g3335375; PIDN:AAC27176.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2G38280; F16M14.21
A:Map position: 2
A:Introns: 64/3; 82/1; 124/3; 184/3; 207/2; 255/1; 275/3; 313/3; 360/3; 411/3; 429/3; 46
C:Superfamily: AMP deaminase

Query Match	84.7%;	Score	2624.5;	DB 2;	Length	600;			
Best Local Similarity	84.5%;	Pred. No.	2.9e-187;						
Matches	475;	Conservative	45;	Mismatches	41;	Indels	1;	Gaps	1;
QY	5	VAPWEKEVINDPCTPKPNPNFTYVPEPKSEHVFTQVDCVHVYADKDKCTESIVPADAT	64						
DB	34	VAPWEKEVINDPCTPKPNPNFTYVPEPKSEHVFTQVDCVHVYADKDKCTESIVPADAT	93						
QY	65	TFPTDLHYILRTVTAAGNTTVCNRLNLEHKKFHLMLNADREFLAQKTAAPHDRDFYVNR	124						
DB	94	AFPTDLHYILRTVTAAGNTTVCNRLNLEHKKFHLMLNADREFLAQKTAAPHDRDFYVNR	153						
QY	125	KVDTHVHHSACMNQKHLRFIKSKLRKEPDEVVIFRDGTMTLKEVFESLDLTGYDLNVD	184						
DB	154	KVDTHVHHSACMNQKHLRFIKSKLRKEPDEVVIFRDGTMTLKEVFESLDLTGYDLNVD	213						
QY	185	LLDVHADKSTFHRFDKFNKYNPCQSRRLRIFIKQNLIOGRFLAELTKQVFSDLSASK	244						
DB	214	LLDVHADKSTFHRFDKFNKYNPCQSRRLRIFIKQNLIOGRFLAELTKQVFSDLSASK	273						

QY 96 KFKFHLMLNADREFLAQKTAHPRDFYNNVRKVDTHVHSSACMNQKHLRLRIKSKLRKEPDE 155
 Db 329 RWNLYLLNEYQETSVSXKNPHRDFYNNVRKVDTHVHSSACMNQKHLRLRIKSKLRKDE 388
 QY 156 VVIFRDGTWTLKEVPESLDLGYDLNVLDDVHADKSTFHPRDFKFNKYNPCOGSRLRE 215
 Db 389 KVIKFDGKLLTLDEVPESLHLAGYDLSIDLDMHAHAKOTFHPRDFKFNKYNPCOGSRLRE 448
 QY 216 IFKQDNLIQGRFLAELTKQVSDLSASKYQMAEYRISYGRKQSEWDQASWVNNELH 275
 Db 449 IFKTNNTYIKGTYLADITKQVIFDLENSKYQCEYRISYGRSLDEWDKSLASWIDNKVI 508
 QY 276 SGNVWLVOIPLRYNNYKEMGIVTSFQNLDDNLFVPLFEVTTDPASHPOLHVLQVGL 335
 Db 509 SHNRWLVOIPLRYDIYKGTIGVQSDICKNLFQPLFEVTKNPOSHPKLHVLQVIGF 568
 QY 336 DLVDDSKPERPPTKHPPTPEQWNTNPNFPAFSAFYAYCYANFTLNKLRKSGMTTIKPR 395
 Db 569 DSVDDSKVDRFHRKYPKPSLEWAPONPPYSLYLYSNVASLNQWRKAGFTVLVR 628
 QY 396 PHAGAGVDVHAAFLPFLCHNTSHGINKRSPVQLYYLGOIGLAMSPLSNLSFLDYH 455
 Db 629 PHCAGADPEHLVSAYLLAHGISHGILLRKVPFVQVLYLDQVGIA MSPLSNNAFLTYD 688
 QY 456 RNPFTFFORGLNLSLTDPLQIHLTKPEPLVEEYSIAASLWKLSSCDLCEIARNVSQOS 515
 Db 689 KNPPFRYFKRGLNLSLTDPLQFSYTRPLEEYSVAAQIYKLSNVDMLCEIARNVLSQOS 748
 QY 516 GFSAHKAHWIGKNYFKRGPAGNDIHRNTVPHIRVOFREMWRNEMKLV 564
 Db 749 GHEAQIKHGWIGKDFKSGVGNVDVVRNTNVPDIRINRYDTLSTLELV 797

RESULT 4

T50996
 probable AMP deaminase [imported] - Neurospora crassa
 N:Alternate names: protein B7F18.150
 C:Species: Neurospora crassa
 C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 31-Mar-2003
 C:Accession: T50996
 R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Partmann, B.; Holland, R.; Nyakatura,
 submitted to the Protein Sequence Database, July 2000
 A:Reference number: Z25286

A:Accession: T50996
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-924 <SCH>
 A:Cross-references: EMBL:AL398891; GSPDB:GN00116; NCSP:B7F18.150
 A:Experimental source: BAC clone B7F18; strain OR74A
 C:Genetics:
 A:Gene: NCSP:B7F18.150
 A:Map position: 6
 A:Introns: 795/2
 C:Superfamily: AMP deaminase

Query Match 52.9%; Score 1641; DB 2; Length 924;
 Best Local Similarity 54.8%; Pred. No. 7.6e-114;
 Matches 309; Conservative 93; Mismatches 150; Indels 12; Gaps 2;

QY 10 KEVINDPCTPKPNPNPFTYVPEPKSEHVFTQVDGVIHVYADKDCETESYVPADATT---F 66
 Db 248 EFDVLEDLPLVPGDPDFTYKLE-----NGVQIYSNDQHKHEADQPAIKIPTLREY 298
 QY 67 FTDLHVLVTAAGNTRTVCHNRLNLEHKFKFHLMLNADREFLAQKTAHPRDFYNNVRK 126
 Db 299 YLDLEAILVSSDGSKSAFRRLOYLEKRFDLVLLGEGYDQCKVKVPHRDFYNNVRK 358
 QY 127 DTHVHSSACMNQKHLRLFIKSKLRKEPDEWIFRDGTWTLKEVPESLDLTCYDLNVDLL 186
 Db 359 DTHVHSSACMNQKHLRLFIKSKLRKEPDEWIFRDGTWTLKEVPESLDLTCYDLNVDLL 418
 QY 187 DVHADKSTFHRDKFNKYNPCOGSRLREIFLKQDNLIQGRFLAELTKQVSDLSASKYQ 246

Db 419 DMHAHTDSFHRDFKFNKYNPCOGSRLREIFLKQDNLIQGRFLAELTKQVSDLSASKYQ 478
 QY 247 MAEYRISYGRKQSEWDQASWVNNELHSGNVWLVOIPLRYNNYKEMGIVTSFQNLDD 306
 Db 479 MVEMRVSYIGKSLDEWDKSLASWVNDKLFSENVRWLVOIPLRYNNYKASGGVNTFEIVR 538
 QY 307 NIFVPLFEVTTDPASHPOLHVLQVGLDVLVDDSKPERPPTKHPPTPEQWNTNPNFPAF 366
 Db 539 NIFEPLFEVTKDPSHPKLIHFLQVIGLSDVDDSKVERLFRKFPVPIKWDGQONPPY 598
 QY 367 SYAYVYCVANFTLNKLRKSGMTTIKPRPHAGAGVDVHAAFLPFLCHNTSHGINKRKS 426
 Db 599 SYWIYLYANWASLNHWRKRGFTFVLRPHCGAGDSEHLAAVLAALCCHSISHGULLLRKV 658
 QY 427 PVLQYLYLGOIGLAMSPLSNLSFLDYHRRNPFTFFORGLNLSLTDPLQIHLTKPEPL 486
 Db 659 PVLQVVFYLEKIGIAMSPLSNNAFLAYERNPMHQYFKRGLNLSLTDPLQFATKEPL 718
 QY 487 VEEYSIAASLWKLSSCDLCEIARNVSQOSGFSHALKAHWIGKNYFKRGPAGNDIHRNTVP 546
 Db 719 IEEYAVAAQIYKLSNVDMLCEIARNVSQOSYERSIKEQWLKGNFDPGAQGNVTMVKTNVP 778
 QY 547 HIRVOFREMWRNEMKLVYSDNEI 570
 Db 779 DRBEFRTYTLLEBERGMVRYNAL 802

RESULT 5

T39261
 amp deaminase - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
 C:Accession: T39261
 R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Davis, P.; Churcher, C.M.
 submitted to the EMBL Data Library, August 1999
 A:Reference number: Z21840
 A:Accession: T39261
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-846 <LYN>
 A:Cross-references: EMBL:AL110295; PIDN:CAB53720.1; GSPDB:GN00067; SPDB:SPBC106.04
 A:Experimental source: strain 972h-; cosmid c106
 C:Genetics:
 A:Gene: SPDB:SPBC106.04
 A:Map position: 2
 A:Introns: 13/2; 72/3
 C:Superfamily: AMP deaminase

Query Match 52.2%; Score 1617; DB 2; Length 846;
 Best Local Similarity 57.8%; Pred. No. 4.1e-112;
 Matches 304; Conservative 84; Mismatches 126; Indels 12; Gaps 3;

QY 43 GVIVHYADKDC-----TESIYPVADATTFDLDHVLVTAAGNTRTVCHNRLNLEHKFK 98
 Db 245 GIFQVYENDSAYIAGTFSFH-IPTIRYIIDLEFLLSASSDGPSSFSFRRLQYLEGRWN 303
 QY 99 FHLMLNADREFLAQKTAHPRDFYNNVRKVDTHVHSSACMNQKHLRLFIKSKLRKEPDEWVI 158
 Db 304 MYMLNEYQELADTKVPHRDFYNNVRKVDTHVHSSALANQKHLRLFIKAKLRKCPNEKI 363
 QY 159 FRDGTWTLKEVPESLDLTCYDLNVDLLDVHADKSTFHRDKFNKYNPCOGSRLREIFL 218
 Db 364 WRDCKELTLQEVFDSLKLTSDLSIDTLDMAHTDTHPRDFKFNKYNPCOGSRLRTIFL 423
 QY 219 KQDNLIQGRFLAELTKQVSDLSASKYQMAEYRISYGRKQSEWDQASWVNNELHSGN 278
 Db 424 KTDNDINGRYLAELTKQVSDLTQKQMAEYRISYGRNREEDWDLAAATIDNELFSPN 483
 QY 279 VVWLVOIPLRYNNYKEMGIVTSFQNLDDNLFVPLFEVTTDPASHPOLHVLQVGLDVL 338
 Db 484 VRWLIOVPRLYDYVYKSGIVETFEVVRNVPFELVETKDPTRTHPKLHVLQVIGFDSV 543

Db 547 KSPNDPWTSEONPPYSYLYMYANIMVNLRRERGLSTFLFRPHCGEAGSITHLSA 606

Qy 411 FLCHNISHGINLRKSPVLQYLYLQIGIAMSPLSNLSFLDYHRNPPTFFQRLNVS 470

Db 607 FLTADNISHGLLKKSPVLQYLYLAQIPAMSPLSNLSFLFYSKNPLREFLHKLHVS 666

Qy 471 LSTDPLQIHLKEPLVEYSSTAASLWKLSSCDLCEIARNVSYYQGSFSAKAWHIGNY 530

Db 667 LSTDPMQPHYTKEALMEBYATAAQVWKLSTCDLCEIARNVSVLQSLGSHQOKFLQNY 726

Qy 531 FKRGAGNDIHRNTVPHIRVQFREMWRNEMKLV 564

Db 727 YKEGPEGNDIRKTNVAQIRMAFRYETLCLNELSFL 760

RESULT 8

S68146

AMP deaminase (EC 3.5.4.6), erythrocyte, splice form 1a - human

N:Alternate names: AMP deaminase isoform E

N:Contains: AMP deaminase splice form 1b

C:Species: Homo sapiens (man)

C:Date: 06-Dec-1996 #sequence revision 13-Mar-1997 #text change 20-Jun-2000

C:Accession: S68146; S68148; A45071; S28149; S27955

R:Mahneke-Zizelman, D.K.; Eddy, R.; Shows, T.B.; Sabina, R.L.

Biochim. Biophys. Acta 1306: 75-92, 1996

A:Title: Characterization of the human AMPD3 gene reveals that 5' exon usage is subject to alternative splicing

A:Reference number: S68146; MUID:96201708; PMID:8611627

A:Accession: S68146

A:Molecule type: DNA

A:Residues: 1-776 <MAH>

A:Cross-references: EMBL:U29925

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1995

A>Note: only a small part of the nucleic acid sequence is shown

A>Note: only a small part of the translation is shown

A>Note: splice form 1a

A:Accession: S68148

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 10-776 <MAW>

A:Cross-references: EMBL:U29925

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1995

A>Note: splice form 1b; Met-10 is the initiator

R:Mahneke-Zizelman, D.K.; Sabina, R.L.

J. Biol. Chem. 267: 20866-20877, 1992

A:Title: Cloning of human AMP deaminase isoform E cDNAs. Evidence for a third AMPD gene

A:Reference number: A45071; MUID:93015995; PMID:1400401

A:Accession: A45071

A:Molecule type: mRNA

A:Residues: 1-216 <MA2>

A:Cross-references: EMBL:M84720; NID:g178548; PIDN:AAAS8365.1; PID:g178549

A>Note: splice form 1a (fragment)

A>Note: sequence extracted from NCBI backbone (NCBIP:116076)

A:Accession: B45071

A>Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 10-776 <MA3>

A:Cross-references: GB:M84721; NID:g178550; PIDN:AAA58366.1; PID:g178551

A>Note: sequence extracted from NCBI backbone (NCBIP:116085)

A>Note: splice form 1b; Met-10 is the initiator

R:Yamada, Y.; Goto, H.; Ogasawara, N.

Biochim. Biophys. Acta 1171: 125-128, 1992

A:Title: Cloning and nucleotide sequence of the cDNA encoding human erythrocyte-specific AMP deaminase

A:Reference number: S28149; MUID:93042002; PMID:1420359

A:Accession: S28149

A:Molecule type: mRNA

A:Residues: 10-776 <YAM>

A:Cross-references: GB:D12775; NID:g219456; PIDN:BA002240.1; PID:g219457

A>Note: splice-form 1b; Met-10 is the initiator

C:Genetics:

A:Gene: GDB:AMPD3

A:Cross-references: GDB:136013; OMIM:102772

A:Map position: 11p15-11p15

A:Introns: 8/1; 83/2; 151/3; 206/1; 279/2; 322/3; 387/3; 431/3; 486/2; 528/3; 583/2; 623

C:Superfamily: AMP deaminase

C:Keywords: alternative initiators; alternative splicing; erythrocyte; hydrolase

F:1-776/Product: AMP deaminase splice form 1a #status predicted <LSPL>

F:10-776/Product: AMP deaminase splice form 1b #status predicted <SSPL>

Query Match 50.0%; Score 1549; DB 2; Length 776;

Best Local Similarity 50.9%; Pred. No. 4.1e-107;

Matches 292; Conservative 93; Mismatches 175; Indels 14; Gaps 4;

Qy 1 PRVRVAPWEKEVINDCPTPKENPNPFTYVBPFS-EHVFQTVGVHIVYADKCTE---- 55

Db 193 PRADTAPPEEGLPDHPHPPPLQEDPYCLDDAPPNLDYLVMQGGILFYDNKMLEHOEP 252

Qy 56 SIYPVADATTFETDLHYILRVTAAGNRTVCHNRLNLEHKKFKFLMLNADREFLAQKTA 115

Db 253 HSLPYPDLETYVDMSHILALITDPTKTYCHRRINFLSEKFSLHEMLNEMSEFKELKSN 312

Qy 116 PHRDFYNNRVKVDTHVHHSACWKNQKHLRFKSKLKEPDEVVIFRDGYTMTLKEVFESLD 175

Db 313 PHRDFYNNRVKVDTHIAAACWKNQKHLRFKHTYQTEPDRTVAEKRGRKITLRQVFDGLH 372

Qy 176 LTGYDLNVDLDDVADKSTFHRFDKFNLYNPGOSRLREIFLKQDNLIQGRFLAELFKQ 235

Db 373 MDPTDLTVDSLVDVAGRTFHRFDKFNLYNPGOSRLREIFLKQDNLIQGRFLAELFKQ 432

Qy 236 VFDLSASKYQWAEYRISYGRKSEWDLASWVNNELHSGVNVVWLQVLPRLYNNYKEM 295

Db 433 VARELESKYQVSEPRLSYGRSPPEWPNLAYWFIQHKVSPNMRWIIQVPRYIDFRSK 492

Qy 296 CIVTSFQNLNDFVPLEFVITDPAHQHVLFLQVGLDLVDDESPPRRPTKHM--- 352

Db 493 KLLNFGKMLENIFLPLFKATINFDHRELHLFLKYVTGFDSDVDESK---HSDHMFSD 548

Qy 353 --PPEOMTVNFPNPAFVYVYCVANFLTKLRESKGMTTKRPHAGEAGDVDDHLAAT 410

Db 549 KSPNPDWTSQNPYSYLYMYANIMVNLRRERGLSTFLFRPHCGEAGSITHLSA 608

Qy 411 FLCHNISHGINLRKSPVLQYLYLQIGIAMSPLSNLSFLDYHRNPPTFFQRLNVS 470

Db 609 FLTADNISHGLLKKSPVLQYLYLAQIPAMSPLSNLSFLFYSKNPLREFLHKLHVS 668

Qy 471 LSTDPLQIHLKEPLVEYSSTAASLWKLSSCDLCEIARNVSYYQGSFSAKAWHIGNY 530

Db 669 LSTDPMQPHYTKEALMEBYATAAQVWKLSTCDLCEIARNVSVLQSLGSHQOKFLQNY 728

Qy 531 FKRGAGNDIHRNTVPHIRVQFREMWRNEMKLV 564

Db 729 YKEGPEGNDIRKTNVAQIRMAFRYETLCLNELSFL 762

RESULT 9

A27366

AMP deaminase (EC 3.5.4.6), skeletal muscle - rat

N:Alternate names: adenylic acid deaminase; AMP aminase; myoadenylate deaminase

C:Species: Rattus norvegicus (Norway rat)

C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 10-Sep-1999

C:Accession: A27366; I57509

R:Sabina, R.L.; Marquetant, R.; Desai, N.M.; Kaletka, K.; Holmes, E.W.

J. Biol. Chem. 262: 12397-12400, 1987

A:Title: Cloning and sequence of rat myoadenylate deaminase cDNA. Evidence for tissue-sp

A:Reference number: A27366; MUID:87308255; PMID:3624265

A:Accession: A27366

A:Molecule type: mRNA

A:Residues: 1-747 <SAB>

A:Cross-references: GB:J02811; NID:g202882; PIDN:AA854086.1; PID:g202883

R:Mineo, I.; Clarke, P.R.H.; Sabina, R.L.; Holmes, E.W.

Mol. Cell. Biol. 10: 5271-5278, 1990

A:Title: A novel pathway for alternative splicing: Identification of an RNA intermediate

A:Reference number: I57509; MUID:90377216; PMID:2398891

A:Accession: I57509

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-17 <RES>

A;Cross-references: GB:M58688; NID:G202875; PIDN:AAA40726.1; PID:G554413
C;Genetics:
A;Gene: AMPD1
A;Introns: 8/1; 12/1
C;Superfamily: AMP deaminase
C;Keywords: hydrolase; muscle

Query Match 48.5%; Score 1505; DB 1; Length 747;
Best Local Similarity 51.2%; Pred. No. 7.3e-104; Indels 16; Gaps 5;
Matches 285; Conservative 87; Mismatches 169

QY 19 PKPNPNPT--VPEPKSEHVFTQVDGVIHVYADKCTE-----SIYPVADATTTFFDLHY 72
DB 188 PKKGEDPFRREDLPANLGVHL-KMKGGVIYVPEAAASRDEPKPYENLDDFLDDMNF 246
QY 73 ILAVTAAGNTRTVCHNRLLNLEHFKFHLMLNADREFLAQKTAHRDFNVKRYDTHVH 132
DB 247 LLALIAAGPVKTYTHRLKFLSKFQVHQLNEMDELMKELKNPHRDFYCKRVDTIHA 306
QY 133 SACMNOKHLRLFTKSLRKEPDEWIFRDGTVMYTLKEVPESLDLTGYDLNVLLDVHADK 192
DB 307 AACMNOKHLRLFTKSLRKEPDEWIFRDGTVMYTLKEVPESLDLTGYDLNVLLDVHADK 366
QY 193 STPHRDFKFNLYKPCQSGRLREIFLKQDNLIQGRFLAELTKQVFSDSLSSASKYQMAEYRI 252
DB 367 QTFQRFDFKFNLYKPCQSGRLREIFLKQDNLIQGRFLAELTKQVFSDSLSSASKYQMAEYRI 426
QY 253 SIYGRKQSEWDQLASWIVNNELHSGNVNMLVQIPRLYNYKEMGIIVTSFQNLNLFVPL 312
DB 427 SIYGRSPDEWSKLSWVFNVRNIIYCPNMTWMIQVRIYDVFRSKNPLFPHFGKMLNIFLEV 486
QY 313 FEVTIPASHPOLHVLKQVVGDLVDDESKPERRPTKGM-----PTPEQWNTNVF 367
DB 487 FEATINPQTHPDLSVFLKHTGTGDSVDDSK-----HSGHMFSSKSPKPEWNTNENPSY 542
QY 368 YIAYCYANLFTLNKLRSEKGMTTIKPRPHAGAGDVHDLAATFLCHNISHGINLRKSP 427
DB 543 YIAYCYANLFTLNKLRSEKGMTTIKPRPHAGAGDVHDLAATFLCHNISHGINLRKSP 602
QY 428 VLOYLYLGOIGLAMSPLNSFLDYHRNPFTFFQGRGLNVLSLSTDDPLQHLTKPLV 487
DB 603 VLOYLYLGOIGLAMSPLNSFLDYHRNPFTFFQGRGLNVLSLSTDDPLQHLTKPLV 662
QY 488 EYYSIAASLWKLSSCDLCEIARNVYQSGFSHAKAHWIGKYNFKRGPAGNDIHRTNVPH 547
DB 663 EYYSIAASLWKLSSCDLCEIARNVYQSGFSHAKAHWIGKYNFKRGPAGNDIHRTNVPH 722
QY 548 IRVQFREMIRNEMKLV 564
DB 723 IRMAYRYETWCYELNLI 739

RESULT 10
I39444
AMP deaminase (EC 3.5.4.6) - human
N;Alternate names: myoadenylate deaminase
C;Species: Homo sapiens (man)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 22-Jun-1999
C;Accession: I39444
R;Sabina, R.L.; Fishbein, W.N.; Pezeshkpour, G.; Clarke, P.R.; Holmes, E.W.
Neurology 42, 170-179, 1992
A;Title: Molecular analysis of the myoadenylate deaminase deficiencies.
A;Reference number: I39444; MUID:9211279; PMID:1370861
A;Accession: I39444
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-747 <RES>
A;Cross-references: GB:M60092; NID:G178543; PIDN:AAA57281.1; PID:G178544
C;Genetics:
A;Gene: GDB:AMPD1
A;Cross-references: GDB:113677; OMIM:102770
A;Map position: 1p13-1p13
C;Superfamily: AMP deaminase

C;Keywords: hydrolase

Query Match 48.3%; Score 1497; DB 2; Length 747;
Best Local Similarity 51.4%; Pred. No. 2.9e-103; Indels 22; Gaps 8;
Matches 289; Conservative 89; Mismatches 162

QY 18 TP--KPNENPF--TVPEPKSEHVFTQVDGVIHVY-----ADKCTESI-YPVADATTTFF 67
DB 185 TTPVKKGEDPRTDNLNENLGVHL-KMKDGVVYVYVPEAAVSKDEPKPLPENLD--TFL 241
QY 68 TDLHYLRTVTAAGNTRTVCHNRLLNLEHFKFHLMLNADREFLAQKTAHRDFNVKRYD 127
DB 242 DDMNLLALIAAGPVKTYTHRLKFLSKFQVHQLNEMDELMKELKNPHRDFYCKRV 301
QY 128 THVHSAACMNOKHLRLFTKSLRKEPDEWIFRDGTVMYTLKEVPESLDLTGYDLNVLLD 187
DB 302 THVHSAACMNOKHLRLFTKSLRKEPDEWIFRDGTVMYTLKEVPESLDLTGYDLNVLLD 361
QY 188 VHAKSTPHRDFKFNLYKPCQSGRLREIFLKQDNLIQGRFLAELTKQVFSDSLSSASKYQ 247
DB 362 VHAGRQTFQRFDFKFNLYKPCQSGRLREIFLKQDNLIQGRFLAELTKQVFSDSLSSASKYQ 421
QY 248 AEYRISIVGRKQSEWDQLASWIVNNELHSGNVNMLVQIPRLYNYKEMGIIVTSFQNLN 307
DB 422 AEPRLSIYGRSPDEWSKLSWVFNVRNIIYCPNMTWMIQVRIYDVFRSKNPLFPHFGKML 481
QY 308 IFVLPFVETIDPASHPOLHVLKQVVGDLVDDESKPERRPTKGM-----PTPEQWNTNVF 362
DB 482 IFMPVFEATINPQADPESVFLKHTGTGDSVDDSK-----HSGHMFSSKSPKPEWNTLEK 537
QY 363 NPAFSYAYCYANLFTLNKLRSEKGMTTIKPRPHAGAGDVHDLAATFLCHNISHGIN 422
DB 538 NPSTYIAYCYANLFTLNKLRSEKGMTTIKPRPHAGAGDVHDLAATFLCHNISHGIN 597
QY 423 LRKSPVLYLGOIGLAMSPLNSFLDYHRNPFTFFQGRGLNVLSLSTDDPLQHLTKPLV 482
DB 598 LRKSPVLYLGOIGLAMSPLNSFLDYHRNPFTFFQGRGLNVLSLSTDDPLQHLTKPLV 657
QY 483 KEPLVEEYIAAQVFKLSTCDMCEVARNVLCQGISHEEKVKFLGNDLYEGPAGNDIRR 542
DB 658 KEPLVEEYIAAQVFKLSTCDMCEVARNVLCQGISHEEKVKFLGNDLYEGPAGNDIRR 717
QY 543 TNVPHIRVQFREMIRNEMKLV 564
DB 718 TNVPHIRVQFREMIRNEMKLV 739

RESULT 11
S59996
AMP deaminase (EC 3.5.4.6) 2 isoform L - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 22-Jun-1999
C;Accession: S59996
R;van den Bergh, F.; Sabina, R.L.
Biochem. J. 312, 401-410, 1995
A;Title: Characterization of human AMP deaminase 2 (AMPD2) gene expression reveals alter
A;Reference number: S59996; MUID:96103174; PMID:8526848
A;Accession: S59996
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: mRNA
A;Residues: 1-605 <RNA>
A;Cross-references: EMBL:U16269; NID:G608496; PIDN:AAB06511.1; PID:G608497
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1994
C;Genetics:
A;Introns: 175/1; 222/2; 360/3; 404/3
A;Superfamily: AMP deaminase
C;Keywords: alternative splicing; hydrolase

Query Match 38.3%; Score 1186.5; DB 2; Length 605;
Best Local Similarity 50.6%; Pred. No. 2.6e-80; Indels 19; Gaps 4;
Matches 222; Conservative 67; Mismatches 131

QY 1 PRVRVAPWEKEVINDPCTPKPNPNFTYVPEPKSEHVFTQVDGVIHVY-----DKDCTES 56

```

Db      181 PALEQHPYEH-----CESTWPGDGL-----GLRWGRGVHVYTRRPFDEHCSV 226
QY      57 IYPVADATTFDLDHLYLRTVTAAGNTRTVCHNRLNLEHKKFHLMLNADREFLAQKTAP 116
Db      227 ELPYDLOQFVADVNVLMALIINGPIKSCYRRLOQLVSSKFQMHVLLNEMKELAAQKVP 286
QY      117 HRDFYVNRKVDTHVHSHACNQHLLRFKSKLRKEPDEVVIFRDTYMTLKEVPESLDL 176
Db      287 HRDFYVNRKVDTHHSHACNQHLLRFKSKLRKEPDEVVIFRDTYMTLKEVPESLDL 346
QY      177 TGVLDNLVDLVDHADKSTHRRDKENLKNPCQSRLREIFELKQDNLIOGRFLAELTKOV 236
Db      347 TAYDLSVDLVDHADKSTHRRDKFNKYNPIGESVLRIFIKTNRVSGKFAHIIKEV 406
QY      237 FSDLSASKYQMAEYRISIIYGRKQSEWDQLASMIVNNELHSGNVMLVQIPRLYNNYKEMG 296
Db      407 MSDLBSKYQNAELRLSIYGRSRDEWDKLARWAMVHRVSPNRMVQVPRLFVYRTKG 466
QY      297 IVTSFQNLNDNFVLPFVETIDPASHPQLHVLKQVVGJDLVDDESKPERRP-TXHMPTP 355
Db      467 QLANFOEMLENIFLPFVETVHPASHPELHLFLEHVDGFSVDDESKPENHVFNLESPLP 526
QY      356 EOWTNVFNPAFXYAYCYANLFTLNKLRSGMTTIKPRHAGAGDVVDHLAATFLLCH 415
Db      527 EAWVEDNPPYALYLYTTFANMAMLNHLRRQRGFHTVLRPHCGEAGPIHHLVSAFMLAE 586
QY      416 NISHGINLRKSPVLOYLYY 434
Db      587 NISHGILLRKAPVLOYLYY 605

RESULT 12
S50801
AMP deaminase homolog YJL070c - yeast (Saccharomyces cerevisiae)
N:Alternate names: probable membrane protein YJL070c; protein HRD888; protein J1095
C:Species: Saccharomyces cerevisiae
C>Date: 23-Aug-1995 #sequence revision 08-Sep-1995 #text_change 19-Apr-2002
C:Accession: S50801; S47120; S56844; S57733
R:Vandenbol, M.; Durand, P.; Dion, C.; Portetelle, D.; Hilger, F.
Yeast 11, 57-60, 1995
A:Title: Sequence of a 17.1 kb DNA fragment from chromosome X of Saccharomyces cerevisiae
A:Reference number: S50799; MUID:95282514; PMID:7762302
A:Accession: S50801
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-888 <VAN>
A:Cross-references: EMBL:Z34288; NID:g498992; PIDN:CAA84052.1; PID:g498996
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1994
R:Vandenbol, M.; Durand, P.; Dion, C.; Portetelle, D.; Hilger, F.
submitted to the EMBL Data Library, June 1994
A:Description: Sequence analysis of a 17.1 kb DNA fragment from chromosome X of Sacchar
A:Reference number: S47117
A:Accession: S47120
A:Molecule type: DNA
A:Residues: 1-888 <VAN>
A:Cross-references: EMBL:Z34288; NID:g498992; PIDN:CAA84052.1; PID:g498996
R:Vandenbol, M.; Durand, P.; Portetelle, D.; Hilger, F.
submitted to the Protein Sequence Database, September 1995
A:Reference number: S56835
A:Accession: S56844
A:Molecule type: DNA
A:Residues: 1-888 <POH>
A:Cross-references: EMBL:Z49345; NID:g1008222; PIDN:CAA89362.1; PID:g1008223; MIPS:YJL07
R:Sor, F.J.
submitted to the EMBL Data Library, June 1995
A:Reference number: S57731
A:Accession: S57733
A:Molecule type: DNA
A:Residues: 1-888 <SOR>
A:Cross-references: EMBL:X88851; NID:g895892; PIDN:CAA61309.1; PID:g895895
C:Genetics:
A:Cross-references: SGD:S0003606

```

A:Map position: 10L
C:Keywords: transmembrane protein

Query Match 20.1%; Score 627; DB 2; Length 888;
Best Local Similarity 30.1%; Pred. No. 1.9e-38;
Matches 184; Conservative 94; Mismatches 207; Indels 126; Gaps 21;

```

QY      66 FFTDHYLRYTAAGNTRTVCHNRLNLEHKKFHLMLNADREFLAQKTAPHRDFYVNRK 125
Db      281 FRDQFAYIIELIQSHKFNESRKLSYLLDKFELFOYLLSKKEILLANKVVPYRDFYNSRK 340
QY      126 VDTTHVHSHACNQHLLRFKSKLRKEPDEVVIFRD-----GYMTLKEVP----- 171
Db      341 VDRDLSGGCISQRLSEYIWEKNLEPERIV-YDDPETSRLKSLRDIFFQGCSSNDQPI 399
QY      172 -----ESLDLTGYLDNLVDLVDHADKSTHRRDKFNKYNPCQSRLREIFELKQDNL 222
Db      400 AIGLKLIDDELD---WVRNIYLDIDYHLPNKVAKLVGKEMRF-----YLLAKVLEFDN 451
QY      223 LIQGRFLAEL-TKQVFSDSLASKYQMAEYRIS--IYGRKQSEWDQLASMIVNNELHSGNV 279
Db      452 FIEGEYLAELFIKVIHILEKSKYQLAQVSNVQFYSSGEDWYKKFSQWLLRWKLVSYNI 511
QY      280 VMLVQIPRLY-NVYKEMGIIVTSFQNLNDNIFVPLF-----EVTIDPASHPQ---LHVFLK 330
Db      512 RWNIIQIARIFPKLFKE-NVVSNFQEFDLIFNPLFLEKEQLPIDSSVNTDIIGLOFFLS 570
QY      331 QVVGDLVDDESKPE--RRPTKHMPTPEQWT-NVFNPAFXYAYCYANLFTLNKLRSK 387
Db      571 NVCSMDLVKESDEYWKFEFTDMNCKPFWTAQGDNPVTAHMYIYKSLAKVNLRSQN 630
QY      388 GMTTIKPRPHAGE-----AGDVHDLAATFLLCHNISHGINLRKSP----- 427
Db      631 LQNTITLRNYCSPSSRTSQFGVDLYFTDQVESLVCNLLC-----NGLLQVEPLWDYAT 686
QY      428 VLQIYLYGQIGLAWSPLSNNSL-----PL----- 452
Db      687 MIQYLFYLFQIPILAAPLSSVSLNSQSKSTFLKNKNVLLHDLKQDQETAKINPSRDITV 746
QY      453 ---DYHNRNPPPTFFQGLNVLSTDDPL-QIHLTKEPLVEEYSTAAISLWKLSSCDLCEI 507
Db      747 GEQSYETNPWMKFMKGLKLSLSKSLYNSSTLEPLIEYSVAASIIYLLNPTDLCLE 806
QY      508 ARNSVYQSGFSGHALKAHWIGKN-----YFKGKPGAGND-----IHRTNVPHIRVOFR 553
Db      807 SRTSVLSSGEGYKAHWIGVGVKKAPVFEENVGVGDWYDTAKDTSIKHNVPIMRRYR 866
QY      554 EMIWRNEMKLV 564
Db      867 KETLDQEWNVF 877

```

RESULT 13

S44546

probable membrane protein YBR284w - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein YBR2021

C:Species: Saccharomyces cerevisiae

C>Date: 08-Jun-1994 #sequence revision 09-Sep-1994 #text_change 19-Apr-2002

C:Accession: S44546; S46166; S39140

R:Holmstrom, K.; Brandt, T.; Kallensee, T.

Yeast 10(Suppl.A), S47-S62, 1994

A:Title: The sequence of a 32420 bp segment located on the right arm of chromosome II fr

A:Reference number: S44537; MUID:94378722; PMID:8091861

A:Accession: S44546

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-797 <HOL>

R:Brandt, T.; Christiansen, C.; Holmstrom, K.; Kallensee, T.

submitted to the Protein Sequence Database, August 1994

A:Reference number: S46157

A:Accession: S46166

A:Molecule type: DNA

A;Residues: 1-797 <BRA>
A;Cross-references: EMBL:Z36153; NID:G536731; PID:G536732; MIPS:YBR284w
C;Genetics:
A;Cross-references: SGD:S0000488
C;Map position: 2R
C;Superfamily: AMP deaminase
C;Keywords: transmembrane protein
F:501-517/Domain: transmembrane #status predicted <TM1>
F:627-643/Domain: transmembrane #status predicted <TM2>

Query Match 20.2%; Score 625; DB 2; Length 797;
Best Local Similarity 30.6%; Pred. No. 2.3e-38;
Matches 182; Conservative 86; Mismatches 213; Indels 114; Gaps 21;

QY 66 FFTDLHYILRVTAAGNTRVTCNRLNLLHKKFHLNADREFLAQKTAHPRDFVNVK 125
DB 219 FREDFEWCUKIRDRSLRSEKRLQVLNKPVPFQHLHSKEMRQKVKPHKDFVNCRR 278
QY 126 VDPVHHSACMNQKHLIRIKSKLRKEPDEVV--IFRDGTYMTLKEVFE-SLDLTG--- 178
DB 279 IDLNLGSGFCFQWQLTEFIWTKLKRPDRVIHQAF-NGSHITLSQLFKVNFETGQFFN 337
QY 179 -----YDLNVLLDVHADKTFTHRPDKFNKYNPCGQSRLEIFL 218
DB 338 GLKIIDDSFLEWKVYILAKYHLVNDMEIHTGS-----HGKQLRY-----YLIAKTEL 386
QY 219 KODNLIOGRFLAELTKQ-VFSDLSASKYQMAEYRISI-----YGRKQSE--WDQLASWIV 270
DB 387 EFDNYINGEYLAELLTKFLIKPOESKYQLCQLSVDFQFVLHYVDSVDNVMWVFANWLN 446
QY 271 NNELHSGNVWLVOIPRLXNVYKEMGIVTSFQNLLDNIIFVPLF--EVTIDPASHPOLHVF 328
DB 447 HYNIPSNIRWNRIRISRIPELVHTGKVKNFQVYLNLIKPLFNAENYHLKSLGPILLKF 506
QY 329 LKQVGLDLV--DDESKPERRPKHMPTEQWNT-VFNPAFSYAYVCYANLFTLNKLR 385
DB 507 LQSVSSIDLQIDTDNYIWKNFYAVSCLPKDWTSGGDNPTISQYMYVYVNLTKLNHIQ 566
QY 386 -----SKGTTIKFRPHAGEAGDVHDHAAATFLA-CHNISHGINL 423
DB 567 ALHQNTFTLRSSCSPTSMNRTSOFNTLNFTET-----TEALNNFLACGGFLNAENL 620
QY 424 -RKSPVLQVLYLQIGLAMSPLSN-----NSFLD----YHRNPFPTFFQ 464
DB 621 WNAPPSLVLYLYLSQIPMVVAPLNSIVDSKPTMLQEQAPGLVLEPSKYKKNPFMKFFE 680
QY 465 RGLNVSLSTDDPLQIHL-LTKEPLVEEYSTAASLWKLSSCDLCEIARNSVYQSGFSHALKA 523
DB 681 MGFKISLSSESILYNNSTYKEPIIEEYSVAASIYRLHSADLCCELLNSVITSGFSSTLKN 740
QY 524 HWIG-----KNYFKRGPA-----GNDIHRNTVPHIRVOPREMIWRNEMKLV 564
DB 741 KWLGVSLASHDYEVNTGFGVDKWDCKPNTSLHNVPPIIRROYSRSTLAGEWRLLI 795

RESULT 14
A37056
AMP deaminase (EC 3.5.4.6), brain - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 25-Jan-1991 #sequence_revision 25-Jan-1991 #text_change 22-Jun-1999
C;Accession: A37056
R;Morisaki, T.; Sabina, R.L.; Holmes, E.W.
J. Biol. Chem. 265, 11482-11486, 1990
A;Title: Adenylate deaminase. A multigene family in humans and rats.
A;Reference number: A37056; MUID:90307656; PMID:2365682
A;Accession: A37056
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-88 <MOR>
A;Cross-references: GB:M38126; NID:G202880; PID:AAA40728.1; PID:G202881
C;Superfamily: AMP deaminase
C;Keywords: hydrolase

Query Match 11.3%; Score 350; DB 2; Length 88;
Best Local Similarity 72.7%; Pred. No. 2.8e-19;
Matches 64; Conservative 7; Mismatches 17; Indels 0; Gaps 0;

QY 396 PHAGEAGDVHDHAAATFLCHNISHGINLRKSPVLQVLYLQIGLAMSPLSNLSFLDYH 455
DB 1 PHCGEAGPIHHLVSAFMAENISHGLLRKAPVLYLYLAQIGTAMSPLSNLSFLSYH 60
QY 456 RNPPETFEQRLNVSLSTDDPLQIHLTK 483
DB 61 RNPLPEYLSRGLMVSLSTDDPLQPFHFK 88

RESULT 15
G97269
adenosine deaminase [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
C;Accession: G97269
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cl.
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: G97269
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-334 <KUR>
A;Cross-references: GB:AE001437; PIDN:AAK80946.1; PID:gl5026062; GSPDB:GN00168
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC3005
C;Superfamily: adenosine deaminase

Query Match 4.8%; Score 149; DB 2; Length 334;
Best Local Similarity 27.3%; Pred. No. 0.0017;
Matches 65; Conservative 39; Mismatches 106; Indels 28; Gaps 11;

QY 314 EVTIDPASHPOLHVLKQVV--GLD-LVDDSKPERRPT-----KHMPTEQWNTVNP 364
DB 100 EIRFAPFOHTQDLNENDVVEAALQDGSKLGHSNLLILCSLRHDPV-ERSIDLNL 158
QY 365 AFSYYAYVCYANL-----FTLNKLRESKGM---TTIKFRPHAGEAGDVHDHAAATFLCH 415
DB 159 ANSYNEGVCADVLDAGNESDFPELHKEAFDAYDNGIKITIHAGETGAENILKSIKLH 218
QY 416 --NISHGINLRKS-PVLQVLYLQIGLAMSPLSN--NSLFLDYHRNPFPTFFQRLNV 470
DB 219 ADRIHGIFAVKSEILOYVIE-NQVPLEMCPKSNVDTKAVKNYKNHPFKKYFDLGKVT 277
QY 471 LSTDDPLQIHLTKPEPLVEEYSTAASLWKLSSCDLCEIARNSVYQSGFSHALKAHWIG 528
DB 278 LNTDRNT---VSNVSLVDYLNLANIFDGTIEETKTIVRNGISASFATEEFKVNLLKK 332

Search completed: December 15, 2003, 13:28:40
Job time : 23 secs